

CITED BY APPLICANT.

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets

(11) Numéro de publication:

**0 379 436**  
**A1**

(12)

**DEMANDE DE BREVET EUROPEEN**

(21) Numéro de dépôt: 90400148.4

(51) Int. Cl.<sup>5</sup>: **G06F 13/18**

(22) Date de dépôt: 19.01.90

(30) Priorité: 20.01.89 FR 8900689

(43) Date de publication de la demande:  
25.07.90 Bulletin 90/30(84) Etats contractants désignés:  
DE GB

(71) Demandeur: **CENTRE NATIONAL DE LA  
RECHERCHE SCIENTIFIQUE (CNRS)**  
13 Quai Anatole France  
F-75700 Paris(FR)

Demandeur: **UNIVERSITE PIERRE ET MARIE  
CURIE PARIS VI**  
4, place Jussieu  
F-75252 Paris Cédex 05(FR)

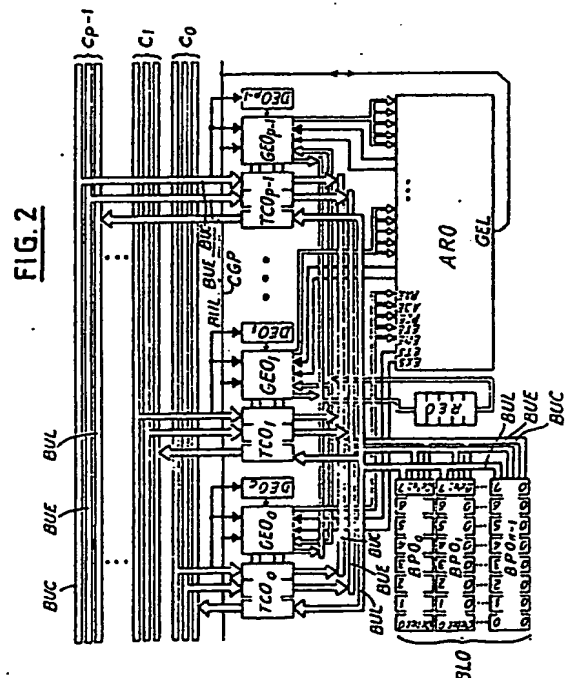
(72) Inventeur: **De Dinechin, Benoît**  
84, Boulevard Raspail  
F-75006 Paris(FR)

(74) Mandataire: **Rodhain, Claude et al**  
Cabinet Claude Rodhain 30, rue la Boétie  
F-75008 Paris(FR)

(54) **Dispositif et procédé d'arbitrage des requêtes et de résolution des conflits liés à l'accès aux mémoires à bancs indépendants pour les machines informatiques.**

(57) L'invention concerne un dispositif et un procédé d'arbitrage destinés aux machines comportant au moins un processeur et une mémoire à bancs logiques (BLO), reliés par des canaux indépendants ( $C_0, C_1 \dots C_{p-1}$ ) fonctionnant simultanément.

Selon l'invention, sur le chemin de chaque banc logique, et pour chaque canal, par un décodeur ( $DEO_0$ ), on vérifie si la requête arrivant est destinée au banc logique, et dans l'affirmative on mémorise dans un tampon ( $TLO_0$ ) géré par un gestionnaire de lecture/écriture ( $GEO_0$ ), les attributs de la requête, puis, dans un arbitreur (ARO) commun à tous les canaux d'accès à ce banc logique, on choisit une requête pouvant passer en mémoire sans déclencher de conflit et on l'envoie à cette mémoire, et en l'absence de requête de priorité absolue non satisfaite, on met à jour les priorités des requêtes en attente, tandis qu'en présence d'une telle requête, on commande le gel du processeur correspondant.



**EP 0 379 436 A1**

## Dispositif et procédé d'arbitrage des requêtes et de résolution des conflits liés à l'accès aux mémoires à bancs indépendants pour les machines informatiques.

L'invention concerne un dispositif et un procédé d'arbitrage des requêtes et de résolution des conflits liés à l'accès aux mémoires à bancs indépendants pour les machines informatiques et plus particulièrement pour les machines parallèles à hautes performances du type mini-supercalculateur ou synthétiseur d'images.

De telles machines comportent un ou plusieurs processeurs connectés sur une mémoire centrale partagée, et l'ensemble de la machine est synchronisé par une horloge unique. Chaque processeur est relié à la mémoire par plusieurs canaux d'accès à celle-ci, indépendants et susceptibles de fonctionner simultanément pour transférer chacun jusqu'à une donnée par cycle de l'horloge.

Il en est ainsi notamment des supercalculateurs industriels, des processeurs de signal, et de la plupart des synthétiseurs d'images et des machines informatiques comportant plusieurs processeurs (le terme processeur désignant ici tout automate apte à gérer, de manière autonome, un flot de requêtes, sur un ou plusieurs canaux mémoire, c'est-à-dire aussi bien des unités de calcul de la machine que des contrôleurs d'entrées-sorties ou encore des gestionnaires de caches). En effet, le volume des données traitées par un supercalculateur ou dans un synthétiseur d'images, impose l'adoption d'une mémoire centrale de grande capacité (parfois supérieure à un giga-octets) ; des considérations de coût et de fiabilité de la machine entraînent, pour les circuits mémoire, l'adoption d'une technologie plus dense et donc plus lente que celle des processeurs, et l'écart de vitesse qui en résulte est aggravé par l'emploi unilatéral (du côté des processeurs) des techniques de pipeline. La seule manière d'approcher le débit idéal d'un mot par cycle et par canal devient alors l'adoption d'une structure de mémoire parallèle. Dans cette configuration, les données sont distribuées dans N mémoires indépendantes appelées bancs physiques. La période de l'horloge étant prise comme unité de temps, la bande passante théorique, en supermots par cycle, d'une mémoire multi-bancs où chaque banc peut servir une requête au plus par  $T_c$  cycles vaut alors  $INF(N/T_c, p)$ ,  $p$  désignant le nombre total de canaux.

Il existe un écart important entre la bande passante théorique d'une mémoire multi-bancs et la bande passante effective. Cet écart impose actuellement un surdimensionnement du nombre de bancs physiques, le facteur  $(N/T_c)/p$  étant de l'ordre de 4 à 6 sur un grand nombre de machines courantes, ce qui limite d'autant le nombre maximal de canaux mémoire et donc le parallélisme

machine réalisable.

L'invention a pour but de remédier à cet inconvénient et ainsi de réduire l'écart en question existant dans les machines à mémoire multi-bancs grâce à une gestion appropriée des accès au système mémoire.

Avant d'exposer d'autres buts que se propose d'atteindre et d'autres problèmes que se propose de résoudre l'invention, rappelons que, dans les systèmes mémoires multi-bancs connus, la plus petite unité de donnée accessible individuellement depuis un processeur, appelée mot, se présente le plus souvent sous la forme d'un octet ; cependant, il est également courant de rencontrer des mots de 32 ou 64 bits ; la quantité maximale que peut transférer un canal, chaque cycle, appelée supermot, correspond, elle, à 64, 128, ou même jusqu'à 512 bits ; la plus petite unité physique de mémorisation organisée par mots est appelée module, et un banc physique est constitué du nombre de modules nécessaire pour former un super-mot à partir de mots ; dans certaines machines, un super-mot a la taille d'un mot, et un banc physique s'identifie alors à un module.

Le nombre total de canaux  $p$ , le nombre de bancs physiques  $N$ , et la durée  $T_c$  des cycles des bancs physiques étant donnés, il existe encore une grande latitude dans l'organisation du système mémoire d'une machine même si l'on se restreint à un type précis d'applications, mais un facteur de choix d'importance se situe au niveau de la connexion entre les bancs physiques et les canaux issus des processeurs, où un compromis doit être atteint entre la complexité, la latence, et le taux de conflits du réseau ; à l'heure actuelle, la meilleure solution semble être la mise en oeuvre de systèmes dits "à barres croisées" complets ou partiels, essentiellement pour des raisons de réduction de latence et de simplicité de contrôle. Ainsi, on connaît des supercalculateurs dans lesquels la mémoire est organisée en quatre sections de huit bancs physiques, offrant chacune une ligne d'accès par processeur ; un réseau à barres croisées  $3 \times 4$  connecte les trois canaux de chaque processeur aux quatre lignes associées d'accès aux sections.

Un deuxième facteur de choix intervient au niveau du règlement des conflits de réseau ou de banc, car la présence de modules incapables de servir plus d'une requête par  $T_c$  cycles, de canaux à fonctionnement simultané et/ou d'un réseau de connexion à capacité limitée, pose le problème de la gestion du système mémoire. Un tel conflit amène à mettre en attente toutes les requêtes conflic-

tuelles sauf une, et la difficulté réside dans le choix d'un algorithme de décision efficace quant au résultat et suffisamment rapide pour suivre le temps de cycle.

Un troisième facteur à prendre en considération, car il influe de manière significative sur les performances d'un système mémoire, est la répartition des adresses dans les bancs, laquelle ne peut généralement être optimisée pour la totalité des applications de la machine. Ainsi, d'après certaines analyses, dans le cas du calcul scientifique, les séquences d'accès sur bancs physiques consécutifs représentent environ 70% des références mémoire, ce qui entraîne l'usage généralisé d'un entrelacement sur la partie basse des adresses. Selon ce schéma, les données d'adresse  $a$  se trouvent dans le banc physique de numéro  $\bar{a}\%N$  ( $\% =$  reste de la division euclidienne), à l'adresse  $a/N$  ( $/ =$  quotient de la division euclidienne).

En conclusion, la mauvaise utilisation de la bande passante d'une mémoire multi-bancs résulte de la combinaison de deux facteurs sur lesquels il n'est possible d'influer que de manière ponctuelle. En effet, à moins de se restreindre à quelques séquences types d'accès, l'apparition de conflits de bancs physiques (et à plus forte raison de conflits de réseaux) s'avère inévitable, ce qui entraîne un retard dans le traitement de certaines requêtes. De plus, il existe toujours des dépendances entre les requêtes présentées au système mémoire (une dépendance entre deux requêtes traduit le fait que la seconde ne peut être envoyée à la mémoire qu'une fois la première acceptée). Les retards pris dans le traitement des requêtes liés par des dépendances se répercutent alors sur le ou les processeur(s) concerné(s) qui cesse(nt) toute émission jusqu'à la résolution des conflits. Il en résulte une sous-utilisation des canaux et donc globalement une mauvaise exploitation de la bande passante du système mémoire.

L'invention a pour but de remédier à ces inconvénients des dispositifs et procédés connus d'arbitrage des requêtes et de résolution des conflits liés à l'accès aux mémoires à bancs indépendants pour les machines informatiques.

A cet effet, il a semblé fructueux de mettre en oeuvre une structure de connexion processeur-mémoire connue et jusqu'ici appliquée seulement aux cas des bancs physiques à un seul module.

En effet, on connaît une structure de connexion processeur-mémoire particulièrement simple et intéressante qui recouvre aussi bien le système à barres croisées complet que le bus partagé, qui repose sur la notion de banc logique. Ce modèle paramétrable appelé L-M (abréviation de Lignes-Modules) étudié de manière relativement complète par F.A Briggs dans le cas des bancs physiques à un module (pas de distinction entre mot et super-

mot), est représenté sur la figure 1. Dans cette structure, un banc logique est constitué par l'association de plusieurs bancs physiques autour d'une ligne constituée d'un bus de contrôle et de deux bus de données réservés respectivement à l'écriture et à la lecture des super-mots ; le partage de la ligne est justifié par le fait que chaque requête ne mobilise un bus que pendant un cycle au plus, alors que les bancs physiques restent limités en débit à une lecture ou à une écriture de super-mot par  $T_c$  cycles ; la désignation "banc logique" provient du fait que ces derniers se comportent, sous réserve d'une bonne répartition des adresses, comme des mémoires pipelinées, capables d'accepter une requête par cycle et de servir les requêtes avec une latence  $T_a$  ( $T_a =$  temps d'accès d'un module ; pour une mémoire statique,  $T_a = T_c$ , alors que pour les mémoires dynamiques,  $T_c = 1,3 T_a$ ). Dans une telle structure L-M, les lignes sont connectées à leur tour aux canaux par l'intermédiaire d'un réseau à barres croisées, ainsi, le chemin d'accès à un banc logique est partagé par tous les processeurs. Ce schéma diffère légèrement de l'organisation du système mémoire d'un certain nombre de machines connues où chaque section offre une ligne d'accès réservée à chaque processeur (dans le cas monoprocesseur, une section équivaut à un banc logique).

Ainsi, l'invention concerne un dispositif d'arbitrage des requêtes et de résolution des conflits liés à l'accès aux mémoires à bancs indépendants des machines informatiques du type comportant au moins un processeur et une mémoire partagée comprenant au moins un banc logique constitué lui-même d'au moins un banc physique, reliés par des canaux d'accès indépendants adaptés à fonctionner simultanément pour transférer jusqu'à une donnée par cycle d'horloge, lesdits canaux comportant chacun un bus d'écriture, un bus de lecture et un bus de contrôle fournissant notamment une information de priorité émise par le processeur, dispositif caractérisé en ce qu'il comporte un contrôleur associé à chaque banc logique pour arbitrer les conflits entre les requêtes d'accès à ce banc logique et, par processeur, un câble de demande de gel pour un signal de demande de gel destiné à interrompre l'émission de requêtes par le processeur, en ce que ce contrôleur comprend, pour chaque canal, un interface comportant des tampons d'entrée d'écriture et de contrôle, et de sortie de lecture, reliés par des bus au canal et au banc logique entre lesquels ils sont insérés, un gestionnaire d'écriture/lecture relié aux tampons pour gérer les transferts entre ceux-ci et d'une part les canaux et d'autre part le banc logique, au bus de contrôle, et au câble de demande de gel, et un décodeur relié au bus de contrôle et au gestionnaire d'écriture/lecture pour indiquer à ce dernier si

des requêtes sont destinées au banc logique, et, associé audit interface, lequel génère une étiquette d'identification de chaque requête à son intention, un arbitreur relié aux gestionnaires d'écriture/lecture et au(x) câble(s) de demande de gel pour sélectionner parmi les requêtes celle qui va accéder au banc logique et demander, si au moins une requête bénéficiant d'une priorité absolue n'est pas satisfaite, le gel du processeur concerné, et un retardeur associé au banc logique et relié aux gestionnaires d'écriture/lecture pour propager des informations de contrôle.

L'invention concerne également un procédé d'arbitrage des requêtes et de résolution des conflits liés à l'accès aux mémoires à bancs indépendants des machines informatiques du type comportant au moins un processeur et une mémoire partagée comprenant au moins un banc logique constitué lui-même d'au moins un banc physique, reliés par des canaux d'accès indépendants adaptés à fonctionner simultanément pour transférer jusqu'à une donnée par cycle d'horloge, lesdits canaux comportant chacun un bus d'écriture, un bus de lecture et un bus de contrôle fournissant notamment une information de priorité initiale émise par le processeur constituée par le nombre de cycles dont on peut retarder la requête, procédé caractérisé en ce que, à chaque arrivée d'une requête par un canal, pour chaque banc logique, on détermine si la requête est destinée audit banc logique, dans l'affirmative, on mémorise provisoirement les paramètres de cette requête, à chaque cycle on choisit parmi toutes les requêtes précédemment mémorisées et n'ayant pas encore accédé au(x) banc(s) physique(s), une requête pouvant passer immédiatement sans provoquer de conflit, et on lui donne accès au banc physique, et en l'absence de requête bénéficiant d'une priorité absolue non satisfaite on décrémente la priorité de toutes les autres requêtes, restées en attente, tandis qu'en présence d'au moins une requête bénéficiant d'une priorité absolue non satisfaite, on donne accès au banc physique à une requête pouvant passer immédiatement sans provoquer de conflit et on envoie un signal de gel provisoire sur le(s) câble(s) de demande de gel relié(s) au(x) processeur(s) à requête(s) de priorité absolue non satisfaite(s) de manière à interrompre l'envoi de nouvelles requêtes par ce(s) processeur(s).

Grâce à ces caractéristiques, l'objectif consistant en la suppression plus ou moins complète des conflits de bancs ou de réseaux, qui s'avère irréaliste dès que l'on quitte les domaines d'application spécialisés, peut être remplacé par celui d'une distribution équitable des requêtes sur l'ensemble des bancs et des lignes en considérant des durées de quelques cycles ; ce dernier critère est plus facile à mettre en oeuvre étant donné qu'il existe

des fonctions matérielles simples et efficaces permettant d'uniformiser de manière satisfaisante la répartition des adresses issues des processeurs sur les modules. De plus, les accumulations locales de requêtes qui subsistent peuvent alors être absorbées en mettant en oeuvre des étages tampons de manière appropriée, sous réserve qu'il existe une certaine indépendance entre les requêtes présentées au système mémoire ; cette dernière condition n'était antérieurement satisfaite que dans le cas des multiprocesseurs où le blocage d'un processeur n'influe pas sur le fonctionnement des autres ; la notion de priorité des requêtes exposée ici représente un progrès considérable dans la mesure où, permettant de spécifier le degré d'indépendance des requêtes, elle autorise l'exploitation des avantages de la mise en oeuvre de tampons au niveau de chaque processeur à un degré inconnu jusqu'ici.

Ainsi, les principes architecturaux adoptés sur les super-calculateurs actuels sont remis en cause par le dispositif et le procédé d'arbitrage des requêtes et de résolution des conflits selon l'invention, particulièrement adaptés aux systèmes mémoires multi-bancs et multi-accès destinés aux machines parallèles hautes performances du type mini-supercalculateur ou synthétiseur d'images.

D'autres caractéristiques et avantages de l'invention ressortiront de la description qui va suivre d'une forme et d'un mode de réalisation préférentiels de l'invention donnés à titre d'exemple non limitatif et illustrés sur les dessins ci-joints dans lesquels :

- la figure 1 montre le schéma synoptique d'une structure L-M connue,
- la figure 2 est le schéma synoptique général d'un dispositif d'arbitrage des requêtes et de résolution des conflits selon l'invention,
- la figure 3 est un schéma synoptique d'un arbitreur équipant le dispositif d'arbitrage de la figure 2,
- la figure 4 est un diagramme temporel d'un étage gestionnaire de canal de l'arbitreur de la figure 3,
- la figure 5 est un schéma synoptique montrant de manière détaillée les circuits d'un électeur de l'arbitreur de la figure 3,
- la figure 6 est un schéma synoptique montrant de manière détaillée les circuits d'une partie d'un bloc interface du dispositif de la figure 2.

Comme on l'a vu, la structure L-M de connexion processeur-mémoire, de base, étudiée par Briggs (figure 1), est destinée à connecter p canaux mémoire C à des bancs logiques BL constitués chacun de N bancs physiques BP. La connexion est effectuée au travers de contrôleurs CO, au moyen de dérivations D et de lignes L à raison d'un contrôleur CO par banc logique BL.

Chaque contrôleur CO est ainsi relié aux canaux C par autant de dérivations D respectives, et à chaque banc physique BP du banc logique BL qui lui est affecté, par une ligne L possédant autant de ramifications qu'il y a de bancs physiques à desservir.

Dans le dispositif selon l'invention, dont une partie seulement est représentée sur la figure 2, la structure est proche de cette structure L-M de base et adaptée au cas où mot et super-mot ne sont pas identiques, un super-mot étant dans l'ensemble choisi constitué de huit octets ; ainsi, un banc physique est constitué de huit modules de chacun un octet, desservis chacun par une sous-ramification. La partie du dispositif selon l'invention représentée sur la figure 2 est celle relative à la connexion des p canaux  $C_0, C_1 \dots C_{p-1}$  issus des divers processeurs, à un seul banc logique BL0 constitué de N bancs physiques  $BP_{00}, BP_{01} \dots BP_{0N-1}$  correspondant chacun à un super-mot de huit mots d'un octet ; selon l'invention, tous les bancs logiques tels que BL0 sont connectés aux p canaux de manière identique au travers d'un contrôleur propre à chaque banc logique ; les canaux  $C_0, C_1 \dots C_{p-1}$  et leur dérivation vers le contrôleur comme les ramifications des lignes vers chaque banc physique et les sous-ramifications, comportent chacun trois bus, à savoir un bus de contrôle BUC, un bus d'écriture BUE, et un bus de lecture BUL. En plus des canaux tels que  $C_0$ , la ligne d'échange entre les processeurs et la mémoire comporte un bus constitué de câbles de demande de gel provisoire des processeurs CGP, et ce bus dessert tous les contrôleurs. Il y a autant de câbles de demande de gel que de processeurs, quoique un seul câble soit représenté sur la figure 2 pour ne pas compliquer le dessin.

Les contrôleurs sont tous constitués de manière identique, et on décrira donc plus précisément le contrôleur relatif à un banc logique BL0. Ce contrôleur comporte principalement un arbitreur AR0 et un retardateur REO uniques, et un certain nombre d'autres étages en autant d'exemplaires qu'il y a de canaux  $C_0, C_1 \dots C_{p-1}$  provenant des processeurs, chaque canal étant affecté de manière exclusive à un processeur. Ces autres étages sont des tampons doubles  $TC_{00}, TC_{01} \dots TC_{0p-1}$ , des décodeurs  $DEO_0, DEO_1 \dots DEO_{p-1}$ , et des étages de gestion de lecture/écriture  $GEO_0, GEO_1 \dots GEO_{p-1}$ . Les tampons  $TC_{01}, TC_{01} \dots TC_{0p-1}$  sont constitués chacun d'un tampon d'entrée (de contrôle et d'écriture) et d'un tampon de sortie (de lecture) reliés chacun (respectivement sur une entrée et une entrée, et une sortie) au bus correspondant du canal qui leur est affecté, au moyen d'un bus de dérivation déjà mentionné, ainsi que (respectivement sur une sortie et une sortie, et une entrée), au bus correspondant de la ligne d'accès

au banc logique. Les décodeurs  $DEO_0, DEO_1 \dots DEO_{p-1}$  ont leur entrée reliée au bus de contrôle de la dérivation. Les étages de gestion de lecture/écriture  $GEO_0, GEO_1 \dots GEO_{p-1}$  ont une entrée reliée à la sortie du décodeur correspondant, une entrée reliée au bus de contrôle de la même dérivation, une entrée reliée au câble CGP de demande de gel du processeur concerné lequel est également relié à l'une des entrées/sorties  $\overline{GEL}$  de l'arbitreur AR0 commune pour tous les canaux affectés au même processeur et délivrant des signaux  $\overline{GEL}$  (ou plus exactement  $\overline{GEL}$ ), une entrée reliée à une sortie EXS correspondante de l'arbitreur, une entrée reliée à une sortie ETS correspondante de l'arbitreur, une entrée reliée à la sortie du retardateur REO, lequel comme l'arbitreur, est commun à tous les étages de gestion du banc logique, et dont l'entrée est reliée à une sortie de l'étage de gestion, et une sortie reliée à des entrées ENE, ETE, PIE, ADE, PAE de l'arbitreur ; chaque étage tampon est par ailleurs également relié à son étage de gestion respectif.

La gestion d'un tel système mémoire L-M arbitré en ce qui concerne les priorités est donc entièrement distribuée au niveau des bancs logiques, les informations partagées se limitant à l'état des lignes de demande de gel reliées aux câbles CGP, et les fonctions des différents étages du contrôleur sont réparties comme suit :

- Chaque bloc interface (il y en a un par canal), comportant un tampon d'entrée, un tampon de sortie, un étage de gestion des écritures dans le tampon d'entrée / lectures dans le tampon de sortie, et un décodeur, se charge d'intercepter sur le canal les requêtes destinées au banc logique dont il fait partie, et génère l'étiquette utilisée par l'arbitreur pour identifier chaque requête.
- L'arbitreur proprement dit, outre les fonctions de gestion des priorités et d'élection des requêtes décrites précédemment, fournit les adresses de lecture provenant de sa sortie ETS aux tampons d'entrée ainsi que les signaux de validation lecture issus de sa sortie EXS, lors de l'envoi vers les bancs physiques des paramètres de la requête élue.
- Le retardateur propage des informations de contrôle correspondant aux adresses d'écriture et aux signaux de validation utilisés pour stocker dans les tampons de sortie les données retournées par les lectures mémoire.

Le processus selon l'invention de mise en oeuvre du contrôleur qui vient d'être décrit fait appel à une notion que l'on peut définir comme étant une "priorité initiale" d'une requête, correspondant au nombre maximal de cycles dont le traitement de la requête peut être retardé par le système mémoire sans faire intervenir de gel provisoire du processeur. Cette priorité initiale est une information de

contrôle prévue pour être émise par le processeur au même titre que l'adresse ou le signal de lecture/écriture ; elle est transmise par le bus de contrôle BUC. Lors d'une lecture, la donnée en sortie des bancs physiques est re-dirigée vers le tampon de sortie associé au canal d'où elle provient, et elle y est inscrite à une adresse égale à la somme de la date d'arrivée et de la priorité initiale de sa requête (modulo une constante d) ; l'adresse de lecture du tampon de sortie est maintenue par un compteur modulo d qui est incrémenté à la fin de chaque cycle en l'absence de gel processeur. De cette manière, la donnée correspondant à une lecture arrive au processeur exactement ( $\text{Min\_L} + \text{priorité initiale}$ ) cycles après l'émission de la requête correspondante, quelque soient la date exacte de passage sur les bancs physiques et le nombre de gels intervenus. Le nombre  $\text{Min\_L}$  représente la latence minimale du système mémoire (a priori  $T_c + 3$ ), et les cycles sont comptés du point de vue du processeur qui, par définition, ne voit pas les gels.

La priorité "courante" d'une requête correspond à sa priorité initiale diminuée du nombre de cycles écoulés depuis son émission par un processeur, sans compter les cycles où ce processeur était gelé.

Grâce à cette mise en oeuvre, en pratique, seule une part réduite des requêtes nécessite un temps de réponse proche du minimum autorisé par le système mémoire. Par exemple, lors du chargement d'un tampon instructions ou d'un registre vectoriel, seuls les premiers mots sont urgents (priorité initiale proche de zéro) car l'usage des données est séquentiel. De même, pour la lecture d'une ligne de cache, le mot à l'origine de la faute est plus urgent que les autres. En revanche, la totalité des requêtes issues d'un contrôleur d'entrées-sorties peuvent accepter une priorité initiale élevée, et il en est de même pour la plupart des écritures mémoire issues de l'exécution d'un code optimisé globalement.

Dans l'hypothèse où il n'existe qu'un processeur, le processus de mise en oeuvre du dispositif qui a été décrit est le suivant : les p canaux émettent tous les cycles entre zéro et p requêtes au total. Au niveau de chaque banc logique, les adresses valides sur les canaux sont décodées pour déterminer si la requête est destinée au banc logique en question. Dans l'affirmative, les paramètres de la requête sont mémorisés dans les registres tampons d'entrée dédiés au canal sur le banc logique en attendant d'être envoyés aux bancs physiques. Le circuit arbitreur AR0 du banc logique est également prévenu qu'une ou plusieurs nouvelles requêtes arrivent, et il en mémorise les attributs suivants :

- une étiquette, qui représente l'adresse des attri-

buts de la requête dans le tampon d'entrée,

- une priorité initiale : le nombre de cycles dont on peut retarder la requête,

- des informations dont le numéro du banc physique concerné par la requête,

- la description des mots accédés à l'intérieur du super-mot.

Simultanément, l'arbitreur AR0 sélectionne parmi les requêtes précédemment mises en attente, l'élue qui va accéder à la mémoire. Pour cela, il choisit parmi celles qui peuvent passer immédiatement sans provoquer de conflit de banc physique, la requête de priorité courante la plus faible. En l'absence de demande de gel, les priorités courantes de toutes les requêtes en attente dans les L arbitreurs du système mémoire sont décrémentées, et un nouveau cycle commence. S'il existe, dans un arbitreur quelconque, une requête de priorité absolue, par exemple zéro, qui n'est pas passée en mémoire, le gel processeur est demandé sur la ligne, en OU câblé, qui relie l'arbitreur AR0 au câble de gel CGP, et aucun arbitreur ne décrémenterait les priorités courantes de ses requêtes. Pendant un gel processeur, l'arrivée de nouvelles requêtes est stoppée et les priorités ne sont plus décrémentées, mais les arbitreurs des différents bancs logiques continuent à écouler normalement les requêtes précédemment mises en attente vers les bancs physiques de leur banc logique.

Le dispositif et le procédé qui viennent d'être décrits, peuvent être appliqués dans un grand nombre de cas, et en particulier aussi bien à un multiprocesseur qu'à un monoprocesseur. Dans le cas de l'application spécifique au multiprocesseur, il est nécessaire, comme on l'a vu, d'une part de supposer que chaque canal est réservé à l'un des processeurs, et d'autre part de prévoir une ligne de gel CGP par processeur ; en effet, il faut noter que dans un multiprocesseur, lorsqu'une requête est arrivée à la priorité absolue zéro et qu'elle ne peut accéder à la mémoire, seul le processeur dont elle est originaire doit être temporairement gelé. Dans ces conditions, on peut par exemple appliquer l'invention à des mémoires présentant des caractéristiques proches de celles de calculateurs existant, par exemple composées de  $N = 64$  bancs physiques de temps de cycle  $T_c = 4$ , avec un entrelacement sur la partie basse des adresses, le nombre de canaux branchés sur cette mémoire étant  $p = 8$ , soit un rapport  $N + T_c$  égal à  $2p$  ; étant donnés ces 64 bancs physiques, et comme d'une part, d'après Briggs, le nombre de lignes L doit être supérieur à p qui vaut 8 sur le système considéré, et d'autre part, le temps de cycle d'un banc physique vaut 4, il semble peu intéressant d'en prévoir moins de quatre par banc logique, soit L inférieur ou égal à 16 ; ainsi, en prenant pour L une puissance de deux, le seul choix possible est

le nombre de seize.

En ce qui concerne l'affectation des bancs physiques aux bancs logiques, on peut assigner les bancs physiques 0, 1, 2, 3 au banc logique 0, les bancs physiques 4, 5, 6, 7 au banc logique 1, ... les bancs physiques 28, 29, 30, 31 au banc logique 15, ou encore les bancs physiques 0, 16, 32, 48 au banc logique 0, les bancs physiques 1, 17, 33, 49, au banc logique 2, ... les bancs physiques 15, 31, 47, 63 au banc logique 15. Ce dernier schéma est le plus performant dans l'absolu, puisque les requêtes sont distribuées plus uniformément sur l'ensemble des bancs logiques dans le cas important des incréments d'adresses de un supermot ; si ce système n'a pas été retenu sur un certain nombre de machines connues, c'est en raison de l'apparition des "conflits liés" qui résultent d'un nombre de lignes égal au temps de cycle des bancs physiques ; le système des priorités selon l'invention, ayant été conçu pour absorber ce genre de difficultés, l'adoption de cette répartition est possible.

L'expérience montre ainsi que si une mémoire telle que celle qui a été définie plus haut est soumise à un flot de requêtes réparties aléatoirement dans l'espace d'adressage, au débit maximal autorisé, par exemple pour huit canaux, les résultats ne sont pas influencés de manière significative par une modification de l'assignation des bancs physiques aux bancs logiques. De plus, les résultats sont très voisins dans le cas d'un octoprocésseur à un canal par processeur, d'un quadriprocésseur à deux canaux par processeur, d'un biprocésseur à quatre canaux par processeur, et d'un monoprocesseur à huit canaux dès que les priorités initiales dépassent huit. Par ailleurs, le monoprocesseur synchrone à huit canaux donne de meilleurs résultats, dès la priorité initiale 4, que l'octoprocésseur aux degrés 0 et 1 de tamponnage, et d'autre part toutes les machines se valent à 5 % près au-delà de la priorité initiale 12 ; cela signifie que l'option multiprocésseur ne constitue, grâce à l'invention, plus une obligation pour le concepteur de machines parallèles dès que la latence dans le traitement des requêtes devient secondaire devant le débit asymptotique de traitement.

La dernière condition se trouve vérifiée en particulier dans le cadre du calcul vectoriel.

Ainsi, l'invention trouve également une application particulièrement intéressante dans le cas d'un monoprocesseur à huit canaux dans lequel les adresses de requêtes sont régulièrement espacées en mémoire ; dans ce cas du calcul vectoriel, environ 90 % des accès correspondent à de tels flots de requêtes régulièrement espacées d'un incréement de  $s$  surpermots, le cas  $s = 1$  en totalisant à lui tout seul environ 70 %. Dans cette situation, c'est avant tout le nombre total  $n(s)$  de

bancs physiques balayés par le flot de requêtes qui fixe l'efficacité du système mémoire ; pour éviter les dégradations catastrophiques de performances qui affectent tous les systèmes mémoire des supercalculateurs sur les incréments  $s$  multiples de 4, celui-ci peut être équipé d'un dispositif de "biaisage" effectuant une transformation particulière pour chaque adresse a émise par le processeur ; un tel système, reposant sur l'emploi d'un opérateur matériel de division euclidienne rapide par un nombre premier permet de concentrer les problèmes de mauvaise répartition des adresses sur un cas rare, et n'impose dans le cas présent aucune autre adaptation de la répartition des bancs logiques/bancs physiques ni de modification du mécanisme d'arbitrage des requêtes. Dans ces conditions, les conflits n'apparaissent plus que si deux ou plusieurs requêtes demandent simultanément une ligne (conflit de réseau), ou si, pour certains incréments, il existe une accumulation locale de requêtes sur le même banc physique ; l'influence de ces conflits disparaît alors dès que l'on laisse au système mémoire une certaine liberté dans le traitement des requêtes (dans cet exemple, priorités initiales supérieures ou égales à 6). Comme on le voit, les problèmes fondamentaux posés par l'exploitation de la bande passante d'une mémoire centrale à bancs indépendants par un monoprocesseur peuvent donc être considérés comme réglés de manière satisfaisante en associant une structure L-M, ce système de gestion des priorités et une fonction simple de biaisage des adresses, ce qui supprime en fin de compte les motivations technologiques pour la réalisation d'un multiprocésseur, puisque l'on obtient alors un système plus simple et plus performant.

Il est également possible de concevoir une machine spécifique permettant de profiter pleinement du dispositif et du procédé selon l'invention.

Une forme de réalisation préférentielle de l'invention pouvant d'ailleurs comporter quelques variantes vis-à-vis du "modèle" décrit plus haut, va donc être maintenant décrite pour l'application particulière à une machine d'une puissance théorique de 160 MIPS ou 80 MFLOPS sur 64 bits pour un temps de cycle de 50 nanosecondes basée sur une technologie TTL et C-MOS courante. La machine en question comporte une unité de traitement à deux multiplieurs flottants pipelinés, deux unités logiques arithmétiques flottantes pipelinées, quatre opérateurs d'adresses et quatre canaux mémoires indépendants ( $p = 4$ ). Le système mémoire de cette machine comporte  $N = 32$  bancs physiques de 8 octets regroupés en  $L = 8$  lignes de  $M = 4$  modules. Le temps d'occupation  $T_c$  des bancs physiques est de deux cycles (inférieur à 200 nanosecondes). Les priorités initiales varient de 1 à 15 ( $d = 16$ ) pour un temps de réponse

minimal de  $T_c + 3$  cycles.

L'architecture de cette machine distingue huit types de données : les octets signés (8 bits), les octets non signés (8 bits), les entiers courts signés (16 bits), les entiers courts non signés (16 bits), les entiers longs (32 bits), les flottants simple précision (32 bits), les flottants double précision (64 bits), et enfin les valeurs logiques (64 bits). Un mot correspond donc à un octet, et un supermot à 8 octets soit 64 bits. Selon les contraintes d'alignement, toute donnée de taille 1 bits doit résider à une adresse multiple de  $2^{1-3}$ . Est par ailleurs introduite la notion d'hyper-mot qui représente une donnée de L super-mots consécutifs où chaque super-mot réside dans un banc logique distinct. La lecture ou l'écriture des hyper-mots s'effectue exactement de la même manière que les lectures ou écritures simples, à la différence que tous les bancs logiques sont sélectionnés lorsqu'une adresse valide se présente sur un canal.

L'exemple de réalisation de l'invention donné fait appel, comme on l'a vu, uniquement à de la technologie TTL ou C-MOS courante, une extension directe des performances pouvant être obtenue par simple changement de technologie (par exemple passage à la technique ECL et au Bi-CMOS).

La partie du dispositif de contrôle selon l'invention qui sera plus précisément l'objet de la description qui va suivre et de la figure 3 est l'arbitreur (par exemple ARO) de ce dispositif. Il faut tout d'abord noter que l'horloge est conçue de manière à pouvoir supporter l'emploi de la logique dynamique sur les chemins critiques. A cette fin, l'horloge délivre deux signaux carrés de fréquence 20 MHz décalés d'un quart de période (12,5 ns), notés dans la suite MCK et SCK ; l'origine des temps est fixée sur la stabilisation du signal  $\overline{\text{GEL}}$ , lors du flanc montant de MCK.

Cet arbitreur doit cycler en moins de 50 nanosecondes, et dans cette durée, il faut inclure le temps de diffusion/échantillonnage de la demande de gel processeur, soit une dizaine de nanosecondes sur un câble coaxial. L'arbitreur est donc pipeliné, la phase insertion des nouvelles requêtes recouvrant celle de l'élection pour l'accès au banc physique. Il en résulte que les priorités initiales de 0 ne sont pas acceptées, bien que des priorités courantes puissent atteindre en interne cette valeur. Le circuit arbitreur, représenté de manière générale sur la figure 3, est constitué de trois types de blocs, à savoir quatre ordonnanceurs  $\text{ORO}_0$ ,  $\text{ORO}_1$ ...  $\text{ORO}_3$  qui correspondent respectivement aux quatre canaux de la machine, et dont le rôle consiste à mémoriser les requêtes incidentes dans les classes qui correspondent à leur priorité initiale, à rechercher la requête de priorité courante la plus faible, et à faire évoluer les priorités selon l'état de

la broche GEL, quatre estimateurs  $\text{ESO}_0$ ,  $\text{ESO}_1$ ... $\text{ESO}_3$  chargés de vérifier si les requêtes proposées par les ordonnanceurs peuvent passer en mémoire sans provoquer de conflit et dans lesquels est disponible une représentation de l'état courant d'occupation des bancs, les mises à jour étant effectuées à la fin de chaque cycle pour tenir compte du passage du temps et de l'envoi éventuel d'une nouvelle requête vers un banc physique, et un électeur ELO qui collecte les requêtes proposées par les ordonnanceurs après leur passage par l'estimateur, qui choisit la requête non conflictuelle de priorité courante la plus faible et en renvoie les caractéristiques aux estimateurs pour la mise à jour des états internes, et également qui active la demande de gel lorsqu'il reste des requêtes à priorité courante 0 qui ne peuvent pas passer en mémoire.

Plus précisément, toutes les entrées de l'arbitreur correspondent à des entrées des ordonnanceurs, de telle sorte que les informations soient transférées directement pour chaque canal à l'ordonnanceur correspondant.

Ainsi, les ordonnanceurs comportent chacun :

- une entrée de validation ENE reliée à l'étage de gestion de lecture/écriture correspondant pour en recevoir un signal de validation indiquant que la requête doit être enregistrée,
- une entrée d'étiquette d'identification ETE reliée également à l'étage de gestion, pour en recevoir un signal d'étiquette d'identification comprise entre 0 et 15 (0 et d-1) sur quatre bits,
- une entrée de priorité initiale PIE reliée à l'étage de gestion pour en recevoir un signal de priorité initiale comprise entre 1 et 15 (1 à d-1) sur quatre bits,
- une entrée de numéro de banc physique ADE reliée à l'étage de gestion pour en recevoir un signal de numéro du banc physique concerné par la requête incidente,
- une entrée de description PAE reliée à l'étage de gestion pour en recevoir un signal de description encodée des octets accédés dans le banc physique.

Dans le cas présent, le numéro de banc physique ne nécessite que deux bits, et la description des octets accédés dans le banc, quatre bits, car il n'existe que quinze possibilités d'occupation des octets d'un supermot de 64 bits si l'on se restreint à la lecture de mots de 8, 16, 32 et 64 bits alignés sur des adresses respectivement multiples de 1, 2, 4 et 8.

En revanche, les ordonnanceurs comportent deux autres entrées qui ne sont pas connectées à une sortie d'un interface canal, mais à une sortie de l'électeur correspondant à l'ordonnanceur pour la première et à la ligne de gel système pour la seconde. Ainsi, les ordonnanceurs comportent :



- une entrée d'extraction EXE recevant un signal d'extraction indiquant que la requête proposée au cycle précédent passe en mémoire, et que toutes les informations la concernant peuvent être retirées,

- une entrée d'autorisation de décrémentation PUE pour autoriser la décrémentation de toutes les priorités des requêtes mises en attente, un gel n'étant pas intervenu.

Mise à part la sortie d'étiquette ETS reliée comme on l'a vu à l'étage de gestion de lecture/écriture correspondant, les sorties des ordonnanceurs sont toutes réservées à l'usage interne de l'arbitreur. Elles sont utilisées soit directement par l'estimateur correspondant et l'électeur, soit indirectement par l'électeur après un regroupement avec la sortie de l'estimateur. Ainsi, les ordonnanceurs comportent également :

- une sortie d'indication de priorité courante PRS, dont le signal de sortie indique la priorité courante de la requête proposée, cette information, entre 0 et 14 (0 et d-2) servant uniquement dans le processus d'élection,

- une sortie de juxtaposition des "champs" des signaux de description des octets accédés et de numéro de banc physique USS, dont le signal de sortie sur six bits correspondant à la requête proposée par l'ordonnanceur est appliqué directement à l'estimateur concerné et à l'électeur respectivement sur des entrées d'identification IDE et USE de ceux-ci et est mémorisé pendant un demi-cycle dans l'électeur pour un usage ultérieur,

- une sortie d'indication de vacuité VAS, fournissant un signal de vacuité indiquant si l'ordonnanceur est vide de telle sorte que les signaux issus des sorties précédentes ETS, PRS et USS ne soient pas pris en compte,

- une sortie de demande de gel DGS pour fournir un signal destiné à l'électeur et formulant une demande de gel par une mise à l'état "1" lorsqu'il existe plusieurs requêtes de priorité zéro dans l'ordonnanceur,

- une sortie d'accélération de gel AGS pour fournir un signal à l'état "1" accélérant le calcul de la requête de gel lorsqu'une requête valide de priorité 0 est présentée par l'ordonnanceur ; l'état de ce signal se retrouve d'ailleurs simplement à partir de celui des signaux de sortie sur les sorties d'indication de priorité courante PRS et d'indication de vacuité VAS.

Les estimateurs comprennent d'une part, outre l'entrée d'identification IDE qui a déjà été mentionnée, sur six bits, pour identifier au moyen du signal de sortie de juxtaposition déjà mentionné, les octets accédés dans le banc physique et le numéro de ce dernier pour la requête proposée par l'ordonnanceur,

- une entrée de paramétrage COE fournissant à

l'estimateur un signal de paramétrage qui lui indique le temps d'occupation en cycles des modules, l'état des broches de l'entrée COE étant câblée de manière fixe,

5 - une sortie de constatation de conflit CCS délivrant un signal positionné à "1" lorsque la requête présentée sur l'entrée d'identification de conflit IDE ne peut passer en mémoire sans provoquer de conflit.

Les estimateurs comprennent également, d'autre part, une entrée de mise à jour MJE pour recevoir un signal de mise à jour de leur représentation de l'état d'occupation des bancs physiques, reliée à cet effet à une sortie MJS de l'arbitreur.

Enfin, ils comportent une entrée de commutation de mode ECE connectée pour recevoir le signal MCK'. SCK'.

De la sorte, les estimateurs fonctionnent selon deux modes commutés par l'entrée de commutation ECE ; lors de la première phase, les caractéristiques de la requête proposée par l'ordonnanceur sont utilisées pour détecter si un conflit de module risque d'intervenir, et sont mis en oeuvre les signaux sur l'entrée d'identification IDE, l'entrée de paramétrage COS, et la sortie de constatation de conflit CCS ; lors de la deuxième phase, les caractéristiques de la requête choisie par l'électeur pour passer en mémoire sont présentées à tous les estimateurs sur leur entrée de mise à jour MJE afin qu'ils mettent à jour leur représentation de l'état d'occupation des bancs physiques.

Les circuits de l'arbitreur nécessitent dans ces conditions, vingt broches d'entrée ou/et de sortie par canal, nombre auquel il faut ajouter :

35 - une borne de gel GEL (sortie de l'électeur) pour des circuits de gel processeur, reliée aux entrées de décrémentation PUE de chaque estimateur et comme on l'a vu au câble de demande de gel processeur CGP, fournissant un signal  $\overline{\text{GEL}}$ ,

40 - une borne de remise à zéro RAZ non représentée sur la figure 3, pour des circuits de remise à zéro créant un signal de remise à zéro RZ,

- deux bornes non représentées, pour les signaux d'horloge MCK et SCK fournis notamment aux estimateurs et aux ordonnanceurs, par une horloge également non représentée,

45 - quatre bornes non représentées, pour les alimentations.

L'électeur comporte également des bornes de sortie d'extraction EXS, en vue de fournir comme on l'a déjà mentionné, des signaux d'extraction respectifs aux ordonnanceurs et en sortie de l'arbitreur.

L'électeur comporte aussi des entrées d'indication de priorité courante PRE respectivement reliées aux sorties d'indication de priorité courante PRS des ordonnanceurs, et des entrées AGE reliées aux sorties d'accélération de gel AGS des ordonnanceurs. Il comporte également des entrées

CEE et HRE respectivement de validité de requête et d'interruption, reliées aux sorties de vacuité VAS, de demande de gel DGS, et d'accélération de requête AGS, des ordonnanceurs, et aux sorties de constatation de conflit CCS des estimateurs par l'intermédiaire de portes  $P_1, P_2, P_3$  ; il existe, naturellement, un couple de ces entrées par groupe ordonnanceur-estimateur, les portes étant au nombre de trois par groupe ordonnanceur-estimateur : une porte ET  $P_1$  a une entrée reliée à la sortie d'accélération de requête AGS et une entrée reliée à la sortie de constatation de conflit CCS, tandis que sa sortie est reliée à une entrée d'une porte OU  $P_2$  dont une autre entrée est reliée à la sortie de demande de gel DGS et dont la sortie est reliée à l'entrée d'interruption HRE de l'électeur, la porte OU  $P_3$  ayant une entrée reliée à la sortie de constatation de conflit CCS et une entrée reliée à la sortie de vacuité VAS, tandis que sa sortie est reliée à l'entrée de validité de requête CEE de l'électeur.

L'arbitreur est donc constitué de circuits gestionnaires correspondant chacun à un canal, comportant chacun un ordonnanceur, un estimateur, et des portes permettant de former certains signaux destinés à l'électeur, éventuellement intégrés en un seul circuit. Cette configuration permet de s'affranchir des limitations de la technologie VLSI employée, le multiplexage des entrées n'étant pas par ailleurs envisageable pour des raisons de performances ; de plus, elle permet de ne pas se limiter à un nombre de canaux a priori.

Le diagramme temporel d'un gestionnaire de canal selon la configuration qui vient d'être décrite est représenté sur la figure 4, sur laquelle aux premier et deuxième signaux d'horloge MCK et SCK sont adjoints des troisième et quatrième signaux d'horloge MC' et SCK' leur correspondant respectivement avec un retard de 7 ns après entrée/amplification dans le circuit ; en règle générale, un signal prend 4 ns pour pénétrer dans le gestionnaire et 10 ns pour en sortir ; ces signaux d'horloge sont approximativement trapézoïdaux ; par commodité, les signaux apparaissant aux entrées et aux sorties des différents circuits sont notés par les deux premières lettres de la référence de l'entrée ou de la sortie à laquelle ils apparaissent ; ainsi, on voit que le signal de décrémentation PU et le signal de priorité initiale PI apparaissant aux entrées PUE et PIE de l'ordonnanceur sont valides du flanc de montée du premier signal d'horloge MCK au flanc de montée du troisième signal d'horloge MCK', tandis que les signaux de validation EN, d'étiquette d'identification ET, de numéro de banc physique AD, et de description PA, apparaissant aux entrées respectives ENE, ETE, ADE, PAE de l'ordonnanceur sont valides du flanc de descente du premier signal d'horloge au

flanc de descente du troisième, et le signal d'extraction EX qui apparaît à l'entrée EXE de l'ordonnanceur lors du flanc de montée du premier est valide jusqu'au flanc de montée du troisième.

La lecture de l'ordonnanceur se produit lorsque le troisième signal d'horloge MCK' est au niveau haut, et l'écriture entre le flanc de descente du quatrième signal d'horloge SCK' et le flanc de montée du troisième ; la lecture de l'estimateur commence entre le flanc de montée du quatrième et le flanc de descente du troisième, pour se terminer lors du flanc de descente de ce quatrième ; l'écriture de l'estimateur s'effectue entre les flancs de montée des troisième et quatrième signaux d'horloge MCK', SCK'.

Le signal d'accélération de requête de gel AG à la sortie de l'ordonnanceur se stabilise peu après le flanc de montée du quatrième signal d'horloge SCK', tandis qu'il en est de même pour le signal d'étiquette ET, les signaux de juxtaposition des champs des signaux de description et de numéro de banc physique US et d'indication de priorité courante PR également à la sortie de l'ordonnanceur, peu avant le flanc de descente du troisième signal d'horloge MCK', et également de même pour les signaux d'interruption de requête HR et de validité de priorité CE à l'entrée de l'électeur entre les flancs de descente des troisième et quatrième signaux d'horloge MCK', SCK' ; tous ces signaux restent stables au moins jusqu'au flanc de montée du quatrième signal d'horloge SCK'.

Dans l'arbitreur ainsi géré, le circuit électeur, qui est la seule partie réalisée en composants standard, est donc destiné à réaliser d'une part l'élection de la requête destinée à la mémoire, et d'autre part le calcul de la demande de gel au niveau du banc logique.

Comme cela a déjà été mentionné, un gel est demandé au niveau du banc logique si l'un au moins des ordonnanceurs contient une requête valide de priorité nulle qui n'est pas passée en mémoire. En raison du fonctionnement pipeliné de l'arbitreur, on anticipe ce calcul pour que la demande de gel n'intervienne pas trop tard. La condition de gel est alors décomposée en trois parties, la première réalisant l'anticipation demandée :

- Si l'un des ordonnanceurs possède plusieurs requêtes à la priorité zéro (signal DG à un), il est certain qu'un gel devra être déclenché pour le cycle suivant.
- Si plusieurs requêtes de priorité zéro sont proposées simultanément à l'électeur, une seule au plus passera en mémoire d'où également un gel à prévoir.
- Si une requête de priorité zéro est déclarée en conflit potentiel par l'estimateur, un gel devra obligatoirement intervenir puisque cette requête sera retardée.

Le signal de demande de gel GR équivaut à la disjonction des trois conditions précédentes, ce qui s'écrit pour un système mémoire à quatre canaux :

$$GR = DG_0 + DG_1 + DG_2 + DG_3 + AG_0 \cdot AG_1 + AG_0 \cdot AG_2 + AG_0 \cdot AG_3 + AG_1 \cdot AG_2 + AG_1 \cdot AG_3 + AG_2 \cdot AG_3 + AG_0 \cdot CC_0 + AG_1 \cdot CC_1 + AG_2 \cdot CC_2 + AG_3 \cdot CC_3$$

En effet, le polynôme symétrique  $\sigma_k$  de  $n$  variables vaut 1 en algèbre de Boole lorsqu'au moins  $k$  d'entre elles ne sont pas nulles (ici  $k = 2$ ). En notant  $HR_i$  le signal égal à  $DG_i + AG_i \cdot CC_i$ , on obtient :

$$GR = HR_0 + HR_1 + HR_2 + HR_3 + AG_0 \cdot AG_1 + AG_0 \cdot AG_2 + AG_0 \cdot AG_3 + AG_1 \cdot AG_2 + AG_1 \cdot AG_3 + AG_2 \cdot AG_3$$

L'élection de la requête pour l'accès à la mémoire s'effectue de manière simple puisque chaque gestionnaire de canal propose une requête dont la priorité courante est indiquée par le champ PR et la validité par CE. Ce dernier signal est le OU logique du signal VA qui indique si l'ordonnancement est vide, et du signal CC déclenché par l'estimateur lorsqu'une requête proposée se trouve en collision avec l'occupation courante des modules. La requête à élire est, parmi les valides, celle dont la priorité est la plus faible. Le signal EX qui indique à chaque ordonnanceur que sa requête a été choisie s'écrit donc (exemple pour le canal 0) :

$$EX_0 = \overline{CE}_0 \cdot ((CE_0 < CE_1) + (CE_0 = CE_1) \cdot (PR_0 < PR_1)) \cdot ((CE_0 < CE_2) + (CE_0 = CE_2) \cdot (PR_0 < PR_2)) \cdot ((CE_0 < CE_3) + (CE_0 = CE_3) \cdot (PR_0 < PR_3))$$

Sachant que  $(CE_0 < CE_1)$  s'écrit pour des valeurs booléennes  $\overline{CE}_0 \cdot CE_1$ , et d'autre part que  $(CE_0 = CE_1)$  équivaut à  $CE_0 \cdot CE_1 + \overline{CE}_0 \cdot \overline{CE}_1$ , on obtient après simplification :

$$EX_0 = \overline{CE}_0 \cdot (CE_1 + (PR_0 < PR_1)) \cdot (CE_2 + PR_0 < PR_2) \cdot (CE_3 + PR_0 < PR_3)$$

Le calcul de  $(PR_i < PR_j)$  s'effectue en recueillant la retenue sortante de l'opération  $PR_j - PR_i$  en complément à 1, soit en réalité  $\overline{PR}_i + PR_j$ . Cette opération est réalisée ici au moyen d'additionneurs TTL standard de la série AS181 ou de la série AS881. Grâce à ce choix, on peut de plus, avantageusement, installer un mécanisme de préséance circulaire permettant de choisir équitablement entre les canaux lorsque deux ou plusieurs requêtes valides de même priorité sont proposées simultanément. En effet, on remarque d'après l'expression des  $EX_i$  que, dans ce dernier cas, aucun d'eux n'est valide en raison du caractère strict des inégalités. La solution consiste à transformer certaines évaluations de  $(PR_i < PR_j)$  en  $(PR_i \leq PR_j)$ , en utilisant la propriété élémentaire suivante :  $(PR_i \leq PR_j) \Leftrightarrow (\overline{PR}_i + PR_j + 1) \geq 2^4$ . Le terme 1 supplémentaire est introduit dans les additionneurs AS881 comme retenue entrante, et le résultat de la comparaison est donné par la valeur de la rete-

nue sortante.

La préséance circulaire entre les quatre canaux comporte quatre périodes. Lors de la première, le canal 0 est plus prioritaire que le canal 1, lui-même plus prioritaire que le canal 2, etc. Lors de la seconde, une permutation circulaire sur les préséances rend le canal 1 le plus prioritaire, suivi du canal 2, ... jusqu'au canal 0. On obtient donc :

$$\begin{aligned} T0: EX_0 &= \overline{CE}_0 \cdot (CE_1 + (PR_0 \leq PR_1)) \cdot (CE_2 + (PR_0 \leq PR_2)) \cdot (CE_3 + (PR_0 \leq PR_3)) \\ EX_1 &= \overline{CE}_1 \cdot (CE_2 + (PR_1 \leq PR_2)) \cdot (CE_3 + (PR_1 \leq PR_3)) \cdot (CE_0 + (PR_1 < PR_0)) \\ EX_2 &= \overline{CE}_2 \cdot (CE_3 + (PR_2 \leq PR_3)) \cdot (CE_0 + (PR_2 < PR_0)) \cdot (CE_1 + (PR_2 < PR_1)) \\ EX_3 &= \overline{CE}_3 \cdot (CE_0 + (PR_3 < PR_0)) \cdot (CE_2 + (PR_3 < PR_1)) \cdot (CE_2 + (PR_3 < PR_2)) \end{aligned}$$

$$\begin{aligned} T1: EX_0 &= \overline{CE}_0 \cdot (CE_1 + (PR_0 < PR_1)) \cdot (CE_2 + (PR_0 < PR_2)) \cdot (CE_3 + (PR_0 < PR_3)) \\ EX_1 &= \overline{CE}_1 \cdot (CE_2 + (PR_1 \leq PR_2)) \cdot (CE_3 + (PR_1 \leq PR_3)) \cdot (CE_0 + (PR_1 \leq PR_0)) \\ EX_2 &= \overline{CE}_2 \cdot (CE_3 + (PR_2 \leq PR_3)) \cdot (CE_0 + (PR_2 \leq PR_0)) \cdot (CE_1 + (PR_2 < PR_1)) \\ EX_3 &= \overline{CE}_3 \cdot (CE_0 + (PR_3 \leq PR_0)) \cdot (CE_1 + (PR_3 < PR_1)) \cdot (CE_2 + (PR_3 < PR_2)) \end{aligned}$$

$$\begin{aligned} T2: EX_0 &= \overline{CE}_0 \cdot (CE_1 + (PR_0 \leq PR_1)) \cdot (CE_2 + (PR_0 < PR_2)) \cdot (CE_3 + (PR_0 < PR_3)) \\ EX_1 &= \overline{CE}_1 \cdot (CE_2 + (PR_1 < PR_2)) \cdot (CE_3 + (PR_1 < PR_3)) \cdot (CE_0 + (PR_1 < PR_0)) \\ EX_2 &= \overline{CE}_2 \cdot (CE_3 + (PR_2 \leq PR_3)) \cdot (CE_0 + (PR_2 \leq PR_0)) \cdot (CE_1 + (PR_2 \leq PR_1)) \\ EX_3 &= \overline{CE}_3 \cdot (CE_0 + (PR_3 \leq PR_0)) \cdot (CE_1 + (PR_3 \leq PR_1)) \cdot (CE_2 + (PR_3 < PR_2)) \end{aligned}$$

$$\begin{aligned} T3: EX_0 &= \overline{CE}_0 \cdot (CE_1 + (PR_0 \leq PR_1)) \cdot (CE_2 + (PR_0 \leq PR_2)) \cdot (CE_3 + (PR_0 < PR_3)) \\ EX_1 &= \overline{CE}_1 \cdot (CE_2 + (PR_1 \leq PR_2)) \cdot (CE_3 + (PR_1 < PR_3)) \cdot (CE_0 + (PR_1 < PR_0)) \\ EX_2 &= \overline{CE}_2 \cdot (CE_3 + (PR_2 < PR_3)) \cdot (CE_0 + (PR_2 < PR_0)) \cdot (CE_1 + (PR_2 < PR_1)) \\ EX_3 &= \overline{CE}_3 \cdot (CE_0 + (PR_3 \leq PR_0)) \cdot (CE_1 + (PR_3 \leq PR_1)) \cdot (CE_2 + (PR_3 \leq PR_2)) \end{aligned}$$

Ce mécanisme se réalise en générant les retenues entrantes des additionneurs au moyen de compteurs circulaires correctement initialisés, et se généralise à un nombre quelconque de canaux sans difficultés. Une préséance fixe entre les canaux serait moins équitable, bien qu'ayant l'avantage de garantir des situations reproductibles.

L'arbitre représenté sur la figure 5 est adapté à ce mode de fonctionnement, et chacun des gestionnaires de canal tels qu'ils ont été définis plus haut (c'est-à-dire constitués chacun d'un ordonnanceur, d'un estimateur et de portes) y est illustré par un bloc unique  $GCO_0, GCO_1 \dots GCO_3$  muni uniquement de ses entrées et sorties autres que les entrées et sorties qui interconnectent ses constituants et que les entrées qui proviennent des étages de gestion de lecture/écriture  $GEO_0, GEO_1 \dots GEO_3$  correspondants. Ainsi, chaque ges-

tionnaire de canal est représenté avec des entrées d'extraction EXE (pour le signal d'extraction EX ou plus exactement  $\overline{EX}$ ), d'autorisation de décrémentation PUE, de mise à jour MJE, et de réarmement RAZ (recevant un signal de réarmement RZ), ainsi que des sorties d'indication de priorité courante PRS, de juxtaposition des champs des signaux de description et de numéro de banc physique USS, d'accélération de calcul de requête de gel AGS, d'interruption HRS (délivrant le signal d'interruption HR ou plus exactement  $\overline{HR}$ ), et de validité de priorité CES (délivrant le signal de validité de priorité CE ou plus exactement  $\overline{CE}$ ).

Les gestionnaires de canal ont, deux à deux, leurs sorties d'accélération de requête de gel AGS reliées à deux entrées respectives d'une porte ET (six portes au total) ; les sorties respectives de ces portes ET, sont deux à deux, reliées à deux entrées respectives d'une porte  $\overline{OU}$  (trois portes au total) ; les sorties des trois portes  $\overline{OU}$ , ainsi que les sorties d'interruption de requête HRS des quatre gestionnaires de canal sont reliées à sept entrées respectives d'une même porte  $\overline{ET}$  dont la sortie fournit le signal de demande de gel GR, lequel est appliqué à l'entrée d'un inverseur constitué par une porte à collecteur ouvert qui, réalisant un OU câblé, fournit en sortie le signal  $\overline{GEL}$  appliqué d'une part aux entrées d'incrément et de décrémentation PUE des gestionnaires de canal et d'autre part à l'entrée d'un registre constitué par une bascule D synchronisée par le deuxième signal d'horloge SCK où il est mémorisé pour fournir en sortie un signal de gel local  $\overline{RG}$  pour les blocs interfaces.

En ce qui concerne la logique d'élection, pour chaque canal, trois comparateurs sont reliés sur une de leurs entrées à la sortie d'indication de priorité courante PRS du gestionnaire du canal correspondant, et sur l'autre entrée à la sortie PRS d'un autre gestionnaire respectif ; les entrées (inverseuses) de retenue des comparateurs sont reliées aux sorties de registres (bascules D) de trois compteurs respectifs comportant chacun quatre registres destinés à réaliser le mécanisme de préséance circulaire ; les sorties (inverseuses) de retenue attaquent chacune une entrée d'une porte ET respective dont une autre entrée est reliée à la sortie de validité de priorité CES dudit autre gestionnaire de canal dont la sortie PRS est reliée à l'entrée du comparateur correspondant ; les trois sorties respectives des trois portes ET correspondant à un même gestionnaire, sont reliées à des entrées d'une porte  $\overline{OU}$ , ainsi que la sortie de validité de priorité du même gestionnaire par un montage approprié.

La sortie de la porte  $\overline{OU}$  (il y en a une par canal) est reliée d'une part à l'entrée d'extraction EXE du gestionnaire correspondant, et à l'entrée

de commande inverseuse d'un amplificateur trois états dont l'entrée principale est reliée à la sortie de juxtaposition des champs des signaux de description et de numéro de banc physique USS du même gestionnaire et dont la sortie est reliée d'une part à l'entrée de mise à jour MJE de tous les questionnaires de l'arbitreur laquelle est par ailleurs reliée à un montage dont le rôle sera mentionné plus loin et connu sous le nom de "pull-down". La sortie de la porte  $\overline{OU}$  est également reliée à une entrée d'une bascule D synchronisée par le premier signal d'horloge MCK, dont la sortie fournit un signal  $\overline{REX}$  qui correspond au signal EX issu de la sortie EXS visible sur la figure 2.

Les compteurs présentent une structure en anneau et leurs bascules sont remises à zéro par le signal RZ et synchronisées par le deuxième signal d'horloge SCK.

Ainsi, en ce qui concerne les logiques de calcul de commande de gel, les signaux d'accélération de requête de gel AG attaquent les ensembles de portes ET- $\overline{OU}$  dont les sorties sont combinées aux signaux d'interruption de requête HR (ou plutôt  $\overline{HR}$ ) par la porte  $\overline{ET}$  à sept entrées ; le signal de demande de gel GR en résultant est ensuite inversé par la porte à collecteur ouvert, tandis que le signal de gel local RG pour les blocs interface est produit comme on l'a vu par mémorisation de l'état du signal de gel  $\overline{GEL}$ .

En ce qui concerne la logique d'élection, pour chaque canal, trois comparaisons sont effectuées avec les priorités proposées par les autres gestionnaires, les retenues entrantes provenant des trois compteurs et les sortantes attaquant un ensemble de portes ET- $\overline{OU}$  qui génère le signal d'extraction ; celui-ci commande la diffusion du signal de description et de numéro de banc physique US de son canal sur le bus de l'entrée du signal de mise à jour MJE, et si aucun signal d'extraction n'est valide, les caractéristiques d'une requête nulle sont forcées par le "pull-down", les sorties de signal d'extraction à destination des blocs interface étant quant à elles mémorisées par la bascule déjà mentionnée.

Comme on l'a vu, l'arbitreur est ainsi la pièce maîtresse du dispositif selon l'invention, mais il est impossible de limiter ce dispositif à cet arbitreur, et il est nécessaire d'intercaler des blocs interfaces sur le chemin des données entre les canaux et les modules ; la composition de ces blocs interfaces a déjà été esquissée, sur un plan très général, et l'on a mentionné que ceux-ci sont constitués chacun de tampons (un tampon d'entrée et un tampon de sortie) et d'un mécanisme de gestion des écritures dans le tampon d'entrée/lectures dans le tampon de sortie, ainsi que d'un décodeur. En fait, un bloc interface est obligatoirement plus complexe, et l'on a représenté sur la figure 6 le détail d'un tel bloc

interface.

Sur cette figure 6, on reconnaîtra un canal  $C_0$ , un banc logique BLO constitué de bancs physiques et du retardateur REO, et un arbitre constitué d'un électeur ELO et de gestionnaires de canal, un seul gestionnaire  $GCO_0$  étant représenté.

Le canal  $C_0$  est relié au banc logique BLO, comme on l'a vu, d'une part au moyen de bus d'écriture et de contrôle reliant en cascade plus précisément le canal  $C_0$  à un démultiplexeur  $DM_1O$ , le démultiplexeur  $DM_1O$  à un tampon d'entrée TEO, le tampon d'entrée TEO à un amplificateur trois états  $SP_1O$  (commandé par le signal REX reçu de l'électeur ELO, et recevant d'autre part un signal Ident destiné à indiquer au banc logique le canal d'origine de la requête), et l'amplificateur trois états  $SP_1O$  au banc logique BLO ; le banc logique BLO est relié d'autre part au canal  $C_0$ , au moyen de bus de lecture, en passant par un tampon de sortie TSO, une logique d'alignement/multiplexage LOO, et un autre amplificateur trois états  $SP_2O$ .

Par ailleurs, le démultiplexeur  $DM_1O$  et le tampon d'entrée TEO sont reliés par l'intermédiaire d'une part d'un décodeur DCO et d'autre part d'un circuit combinatoire PAO à partir de sorties PAS, LOS du démultiplexeur  $DM_1O$  ; ces sorties sont reliées par ailleurs, respectivement, aux entrées de description PAE et de numéro de banc physique ADE du gestionnaire de canal  $GCO_0$  et plus précisément de l'ordonnanceur.

Le canal  $C_0$  est également relié à l'entrée d'un démultiplexeur  $DM_2O$  dont la sortie est reliée à l'entrée d'une porte ET dont la sortie est reliée à l'entrée d'un compteur modulo  $d (= 16)$   $CP_1O$  synchronisé par le premier signal d'horloge MCK et dont la sortie est reliée d'une part, à l'entrée d'étiquette ETE du gestionnaire  $GEO_0$  et d'autre part à une entrée d'adresse d'écriture AEE du tampon d'entrée TEO la sortie de la porte ET est également reliée à l'entrée d'un inverseur dont la sortie est connectée à une entrée de validation d'écriture EVE du tampon d'entrée TEO et à l'entrée de validation ENE du gestionnaire de canal  $GCO_0$  pour lui fournir son signal de validation  $EN$  ; par ailleurs, le tampon d'entrée TEO comporte également une entrée d'adresse de lecture ALE reliée à la sortie d'étiquette ETS du gestionnaire de canal  $GCO_0$  et une entrée de commande de lecture/écriture LEE pour recevoir le deuxième signal d'horloge SCK (de lecture/écriture).

Le banc logique BLO et le canal  $C_0$  sont reliés, en dehors du bus de lecture, qui, comme les bus d'écriture et de contrôle n'a pas besoin d'être repéré et décrit en détail, par des lignes provenant du banc logique et attaquant le tampon de sortie notamment sur des entrées d'adresse d'écriture AEE et de validation d'écriture VEE pour en rece-

voir des signaux AE et  $VE$ .

Le tampon de sortie TSO est par ailleurs synchronisé, sur une entrée LEE de commande de lecture/écriture, par le premier signal d'horloge MCK ou plus exactement  $\overline{MCK}$  (signal lecture/écriture). Il comporte des entrées d'adresse de lecture ALE reliées d'une part à des sorties d'un deuxième compteur modulo  $d$  (avec  $d = 16$ )  $CP_2O$  dont les retenues entrantes sont fournies par le signal de gel local  $\overline{RG}$ . Les sorties du compteur sont reliées également à des entrées d'un premier additionneur dont d'autres entrées sont reliées aux sorties d'un second additionneur dont les entrées pour un premier opérande sont reliées à la sortie de priorité du démultiplexeur  $DM_2O$  et dont les entrées pour un second opérande sont câblées pour représenter un nombre fixé, par exemple  $Mem L + 2$ , selon la forme de réalisation ; la sortie du premier additionneur est reliée à une entrée du tampon d'entrée TEO ; la sortie de priorité du démultiplexeur  $DM_2O$  est également reliée aux entrées du décodeur  $DEO_0$ , décodeur dont des sorties inverseuses attaquent un système de registres à décalage à chargement parallèle (bascules D), par l'intermédiaire de portes ET (sauf le dernier, attaqué par un inverseur) ; la sortie du démultiplexeur  $DM_2O$  est également reliée à l'entrée de priorité initiale PIE du gestionnaire de canal  $GCO_0$ .

Les entrées d'horloge des registres à décalage sont reliées à la sortie d'un système logique constitué par une porte ET et un inverseur mis en cascade, la porte ET ayant une entrée alimentée par le premier signal d'horloge  $\overline{MCK}$  et une autre entrée par le signal de gel local  $\overline{RG}$ . Par ailleurs, la commande de chargement parallèle des registres à décalage est reliée à la sortie d'une porte ET dont une entrée reçoit le signal de sortie de la porte ET disposée en sortie du deuxième démultiplexeur  $DM_2O$  et une autre entrée reçoit le signal de sortie de ce démultiplexeur signifiant que la requête prise en compte correspond à une lecture mémoire ; la sortie de la dernière bascule du registre à décalage commande le second amplificateur trois états  $SP_2O$ .

Ainsi, les blocs interfaces canal se situant sur le chemin des données entre les canaux et les modules permettent d'offrir une capacité de stockage temporaire des attributs des requêtes, réalisant ainsi le découplage temporel entre la date d'émission et celle du passage en mémoire essentiel à l'efficacité du mécanisme de gestion des priorités qui est l'objet de l'invention. Les sorties des deux bancs de registres constituant les deux tampons sont amplifiées pour attaquer le bus, et les signaux de contrôle (validation, adresse) sont générés par ce bloc interface en ce qui concerne l'écriture dans le tampon d'entrée (paramètres des requêtes) et la lecture du tampon de sortie (résultats des lectures

mémoire) ; en revanche, la lecture du tampon d'entrée est commandée par l'arbitreur (adresse ET et validation par REX), et le contrôle de l'écriture dans les tampons de sortie provient de l'élément de délai associé aux bancs physiques.

Lorsqu'une requête est destinée au banc logique dont fait partie le bloc interface canal, la logique de décodage valide l'écriture des paramètres associés dans le tampon d'entrée, et déclenche l'enregistrement des attributs par le gestionnaire. Le signal provenant de la porte ET montée en sortie du démultiplexeur DM<sub>2</sub>O est également injecté comme retenue entrante du premier compteur modulo d (= 16) qui génère les adresses d'écriture dans le tampon d'entrée, lesquelles sont interprétées comme signal d'étiquette des requêtes par le gestionnaire. Le tampon d'entrée mémorise la donnée éventuelle à inscrire en mémoire, l'adresse à envoyer aux modules, un signal de lecture/écriture mémoire (plus exactement lecture/écriture), l'adresse de banc physique entièrement décodée par le décodeur DCO, le signal de description PA transcodé en occupation des octets par le circuit combinatoire PAO, et enfin l'adresse d'écriture dans le tampon de sortie de l'éventuelle donnée en retour d'une lecture mémoire, fournie par le premier additionneur.

Lorsque la requête réalise une lecture mémoire, le fait qu'une donnée doit être diffusée sur le bus de retour du canal au bout d'un nombre de cycle égal à la priorité initiale plus Min\_L (sans compter les gels) est enregistré dans le registre à décalage à chargement parallèle. En mode normal, ce registre décale un mot de contrôle dont le bit sortant commande l'amplificateur de lecture du tampon de sortie. L'adresse de lecture du tampon de sortie est fournie par le deuxième compteur modulo d, incrémenté à chaque cycle en absence de demande de gel. Lorsqu'une requête en lecture est décodée par le bloc interface, le bit de commande associé est entré dans le registre à décalage au moyen d'un dispositif qui effectue le OU logique du mot de contrôle précédent, et de la sortie du décodeur à d sorties qui reçoit en entrée la valeur de la priorité.

L'adresse d'écriture dans le tampon de sortie est calculée en prenant la valeur courante du compteur qui génère les adresses de lecture, et en lui ajoutant la priorité de la requête plus une constante. La valeur de cette dernière est à ajuster en fonction de la latence des modules mémoire (autour de Mem\_L + 2), de même que la longueur de la partie du registre à décalage qui se trouve entre le décodeur et l'amplificateur de lecture.

Bien entendu, l'invention n'est pas limitée aux formes et aux modes de réalisation ci-dessus décrits et représentés, et on pourra prévoir d'autres

formes et modes de réalisation sans sortir de son cadre ; ainsi, dans certaines réalisations, la priorité absolue peut ne pas correspondre à une seule valeur mais à plusieurs, par exemple 0 et 1 ; cette modification s'avère nécessaire en particulier dans le cas des fréquences de fonctionnement élevées, pour lesquelles les demandes de gel au cycle t agissent sur le reste de la machine au plus tôt au cycle t + 2. De plus et indépendamment, il est possible de simplifier la réalisation lorsque le temps de cycle T<sub>c</sub> des bancs physiques vaut 1 ; dans cette hypothèse, chaque gestionnaire de canal, par exemple GCO<sub>0</sub>, est réduit à son ordonnanceur ORO<sub>0</sub>, tandis que les entrées ADE et PAE, la sortie USS, et les portes P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub>, devenues inutiles, sont supprimées.

## Revendications

1 - Dispositif d'arbitrage des requêtes et de résolution des conflits liés à l'accès aux mémoires à bancs indépendants des machines informatiques du type comportant au moins un processeur et une mémoire partagée comprenant au moins un banc logique (BLO) constitué lui-même d'au moins un banc physique (BPO<sub>0</sub>, BPO...BPO<sub>N-1</sub>), reliés par des canaux d'accès indépendants adaptés à fonctionner simultanément pour transférer jusqu'à une donnée par cycle d'horloge, lesdits canaux comportant chacun un bus d'écriture (BUE), un bus de lecture (BUL) et un bus de contrôle (BUC) fournissant notamment une information de priorité émise par le processeur, dispositif caractérisé en ce qu'il comporte un contrôleur associé à chaque banc logique pour arbitrer les conflits entre les requêtes d'accès à ce banc logique et, par processeur, un câble de demande de gel (CGP), pour un signal de demande de gel destiné à interrompre l'émission de requêtes par le processeur, en ce que ce contrôleur comprend, pour chaque canal, une interface comportant des tampons (TCO<sub>0</sub>) d'entrée (TEO) d'écriture et de contrôle, et de sortie (TSO) de lecture, reliés par des bus au canal et au banc logique entre lesquels ils sont insérés, un gestionnaire d'écriture/lecture (GEO<sub>0</sub>) relié aux tampons pour gérer les transferts entre ceux-ci et d'une part les canaux et d'autre part le banc logique, au bus de contrôle, et au câble de demande de gel, et un décodeur (DEO<sub>0</sub>) relié au bus de contrôle et au gestionnaire d'écriture/lecture pour indiquer à ce dernier si des requêtes sont destinées au banc logique, et, associé audit interface, lequel génère une étiquette d'identification de chaque requête à son intention, un arbitreur (ARO) relié aux gestionnaires d'écriture/lecture et au(x) câble(s) de demande de gel pour sélectionner parmi les requêtes celle qui va accéder au banc logique et demander,

si au moins une requête bénéficiant d'une priorité absolue n'est pas satisfaite, le gel du processeur concerné, et un retardeur (REO) associé au banc logique et relié aux gestionnaires d'écriture/lecture pour propager des informations de contrôle.

2 - Dispositif selon la revendication 1, caractérisé en ce que l'arbitreur (ARO) comporte d'une part, un gestionnaire de canal ( $GCO_0, GCO_1 \dots GCO_3$ ) associé à chaque canal, comprenant un ordonnanceur ( $ORO_0, ORO_1 \dots ORO_3$ ) relié au gestionnaire d'écriture/lecture ( $GEO_0, GEO_1 \dots GEO_3$ ) correspondant pour en recevoir les attributs des requêtes arrivant, et d'autre part un électeur (ELO) relié aux gestionnaires de canal pour réaliser l'élection de la requête destinée à la mémoire en fonction des informations provenant du gestionnaire de canal et calculer les demandes de gel.

3 - Dispositif selon la revendication 1, caractérisé en ce que l'arbitreur (ARO) comporte d'une part, un gestionnaire de canal ( $GCO_0, GCO_1 \dots GCO_3$ ) associé à chaque canal, comprenant un ordonnanceur ( $ORO_0, ORO_1 \dots ORO_3$ ) relié au gestionnaire d'écriture/lecture ( $GEO_0, GEO_1 \dots GEO_3$ ) correspondant pour en recevoir les attributs des requêtes arrivant, un estimateur ( $ESO_0, ESO_1 \dots ESO_3$ ) associé à chaque ordonnanceur pour recevoir une représentation de l'état des bancs physiques et détecter les risques de conflit, et des portes ( $P_1, P_2, P_3$ ) dont les entrées sont reliées à des sorties du gestionnaire de canal et de l'estimateur, et d'autre part un électeur (ELO) relié à la sortie desdites portes, aux gestionnaires de canal, et aux estimateurs, pour réaliser l'élection de la requête destinée à la mémoire en fonction des informations provenant de ces éléments du gestionnaire de canal, transmettre des informations de mise à jour aux estimateurs, et calculer les demandes de gel.

4 - Dispositif selon la revendication 2 ou 3, caractérisé en ce que l'arbitreur (ARO) comporte des entrées (ENE, ETE, PIE, ADE, PAE) reliées à des sorties des gestionnaires d'écriture/lecture ( $GEO_0, GEO_1 \dots GEO_3$ ) pour en recevoir respectivement des signaux de validation, d'étiquette d'identification, de priorité initiale, de numéro du banc physique concerné par la requête, et, en présence d'estimateurs, de description encodée des octets accédés dans le banc physique, ainsi que des sorties (EXS, ETS) reliées également aux gestionnaires d'écriture/lecture pour leur fournir respectivement des signaux de validation lecture et d'adresse de lecture, et une borne de gel reliée au câble de demande de gel (CGP).

5 - Dispositif selon la revendication 3, caractérisé en ce que l'électeur (ELO) comporte un circuit de création d'un signal de demande de gel (GEL) comportant des portes dont les entrées sont reliées à des sorties (AGS) correspondantes de gestionnaires de canal ( $GCO_0, GCO_1 \dots GCO_3$ ), deux à

deux, pour en recevoir des signaux d'accélération de calcul de requête de gel, et un circuit de création d'un signal de gel local (RG) pour les blocs interface, ainsi qu'une logique d'élection comportant, pour chaque canal ( $C_0, C_1 \dots C_3$ ), des comparateurs reliés à des sorties (PRS) correspondantes des gestionnaires de canal, deux à deux, pour en recevoir des signaux d'indication de priorité courante, et un ensemble de portes dont certaines ont leurs entrées reliées respectivement aux sorties de ces comparateurs et à des sorties (CES) des gestionnaires de canal pour en recevoir un signal de validité de priorité, en vue de créer un signal de validation (REX) et un signal d'extraction indiquant que la requête proposée peut passer en mémoire et destiné à être appliqué à une entrée d'extraction (EXE) du gestionnaire de canal.

6 - Dispositif selon la revendication 5, caractérisé en ce que l'électeur (ELO) comporte des compteurs constitués de registres connectés circulairement et reliés chacun à une entrée de retenue d'un comparateur.

7 - Dispositif selon la revendication 1, caractérisé en ce que chaque canal ( $C_0$ ) est relié à un banc logique (BLO) au moyen du bus d'écriture et du bus de contrôle, reliant en cascade le canal à un démultiplexeur ( $DM_1O$ ), le démultiplexeur au tampon d'entrée (TEO), le tampon d'entrée à un amplificateur trois états ( $SP_1O$ ), et l'amplificateur trois états au banc logique, et au moyen du bus de lecture, reliant en cascade le banc logique au tampon de sortie (TSO), le tampon de sortie à une logique d'alignement/multiplexage ( $LOO$ ), la logique d'alignement/multiplexage à un autre amplificateur trois états ( $SP_2O$ ), et l'amplificateur trois états au canal.

8 - Dispositif selon la revendication 7, caractérisé en ce que le décodeur ( $DEO_0$ ) a ses entrées reliées à un démultiplexeur ( $DM_2O$ ) relié au canal ( $C_0$ ) et ses sorties reliées à un système de registres à décalage à chargement parallèle destiné à décaler un mot de contrôle commandant un amplificateur de lecture dans le tampon de sortie (TSO).

9 - Procédé d'arbitrage des requêtes et de résolution des conflits liés à l'accès aux mémoires à bancs indépendants des machines informatiques du type comportant au moins un processeur et une mémoire partagée comprenant au moins un banc logique (BLO) constitué lui-même d'au moins un banc physique ( $BPO_0, BPO_1 \dots BPO_{N-1}$ ), reliés par des canaux d'accès ( $C_0, C_1 \dots C_{p-1}$ ) indépendants adaptés à fonctionner simultanément pour transférer jusqu'à une donnée par cycle d'horloge, lesdits canaux comportant chacun un bus d'écriture (BUC), un bus de lecture et un bus de contrôle fournissant notamment une information de priorité initiale émise par le processeur, constitué par le nombre de cycles dont on peut retarder la requête,

procédé caractérisé en ce que, à chaque arrivée d'une requête par un canal, pour chaque banc logique, on détermine si la requête est destinée audit banc logique, dans l'affirmative, on mémorise provisoirement les paramètres de cette requête, à chaque cycle on choisit parmi toutes les requêtes précédemment mémorisées et n'ayant pas encore accédé au(x) banc(s) physique(s), une requête pouvant passer immédiatement sans provoquer de conflit et on lui donne accès au banc physique, et, en l'absence de requête bénéficiant d'une priorité absolue non satisfaite on décrémente la priorité de toutes les autres requêtes, restées en attente, tandis qu'en présence d'au moins une requête bénéficiant d'une priorité absolue non satisfaite, on envoie un signal de gel provisoire (GEL) sur le(s) câble(s) de demande de gel relié(s) au(x) processeur(s) à requête(s) de priorité absolue non satisfaite(s) de manière à interrompre l'envoi de nouvelles requêtes par ce(s) processeur(s).

10 - Procédé selon la revendication 9, caractérisé en ce que, lors du gel provisoire du processeur, on donne accès au banc physique à une requête pouvant passer immédiatement sans provoquer de conflit.

11 - Procédé selon l'une quelconque des revendications 9 et 10, caractérisé en ce que, pour arbitrer l'accès à la mémoire, on sélectionne, pour chaque ordonnanceur (ORO<sub>0</sub>, ORO<sub>1</sub>, ORO<sub>2</sub>, ORO<sub>3</sub>), la requête de priorité courante la plus faible, on étiquette cette requête avec un signal de validité (CE), et on choisit parmi les requêtes étiquetées, la requête valide de priorité la plus faible.

12 - Procédé selon l'une quelconque des revendications 9 et 10, caractérisé en ce que si une requête bénéficiant d'une priorité absolue n'a pas été satisfaite, on interrompt la décrémentation des requêtes correspondant à tous les bancs logiques (BLO), mais on continue à écouler normalement les requêtes en attente vers les bancs physiques (BPO<sub>0</sub>, BPO<sub>1</sub> ... BPO<sub>N-1</sub>).



FIG. 1

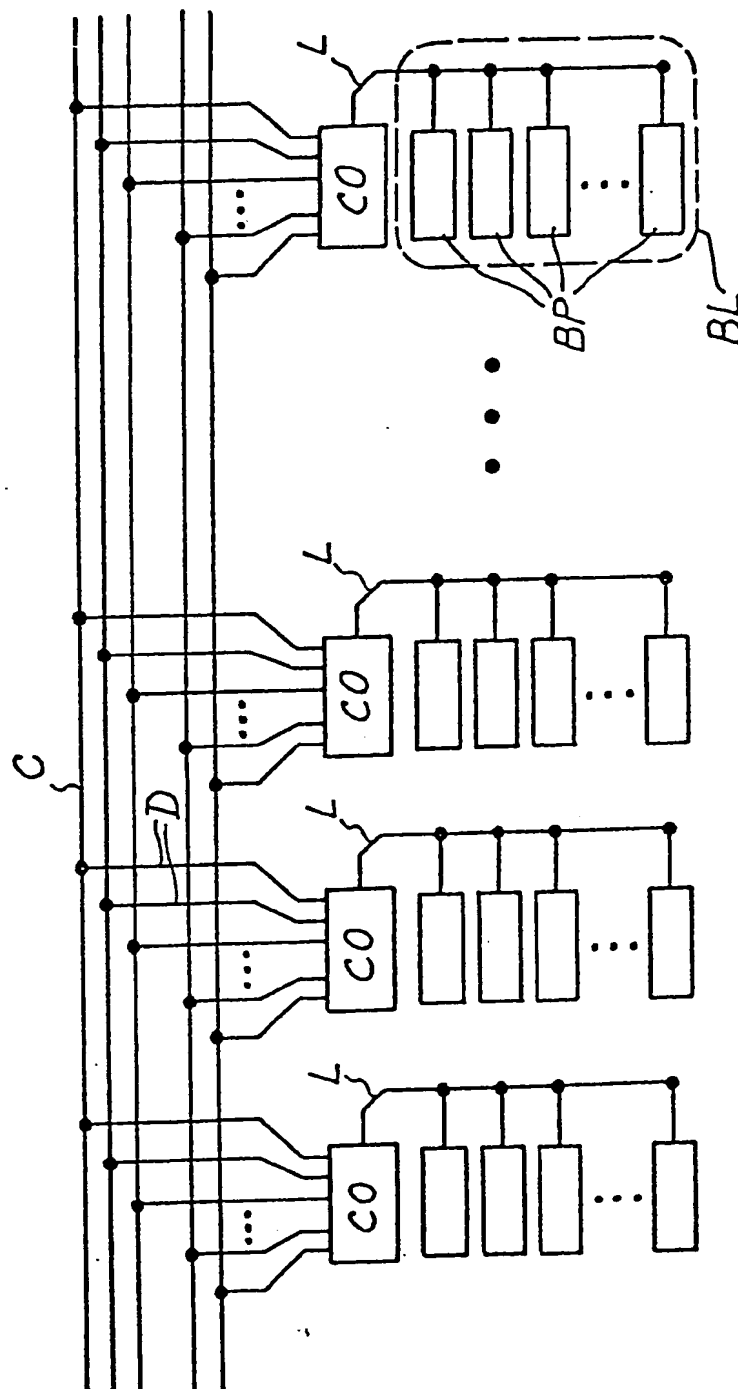
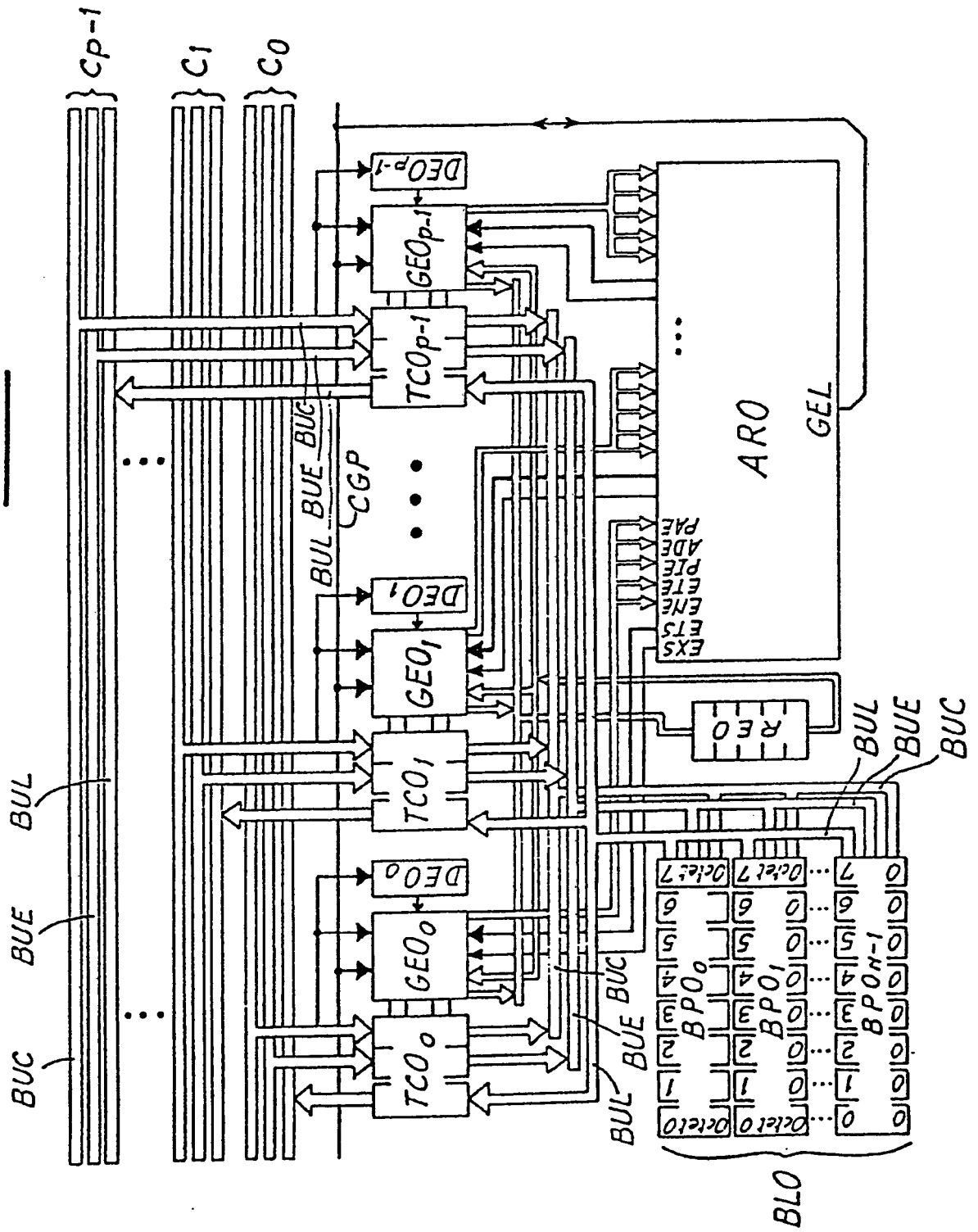


FIG. 2



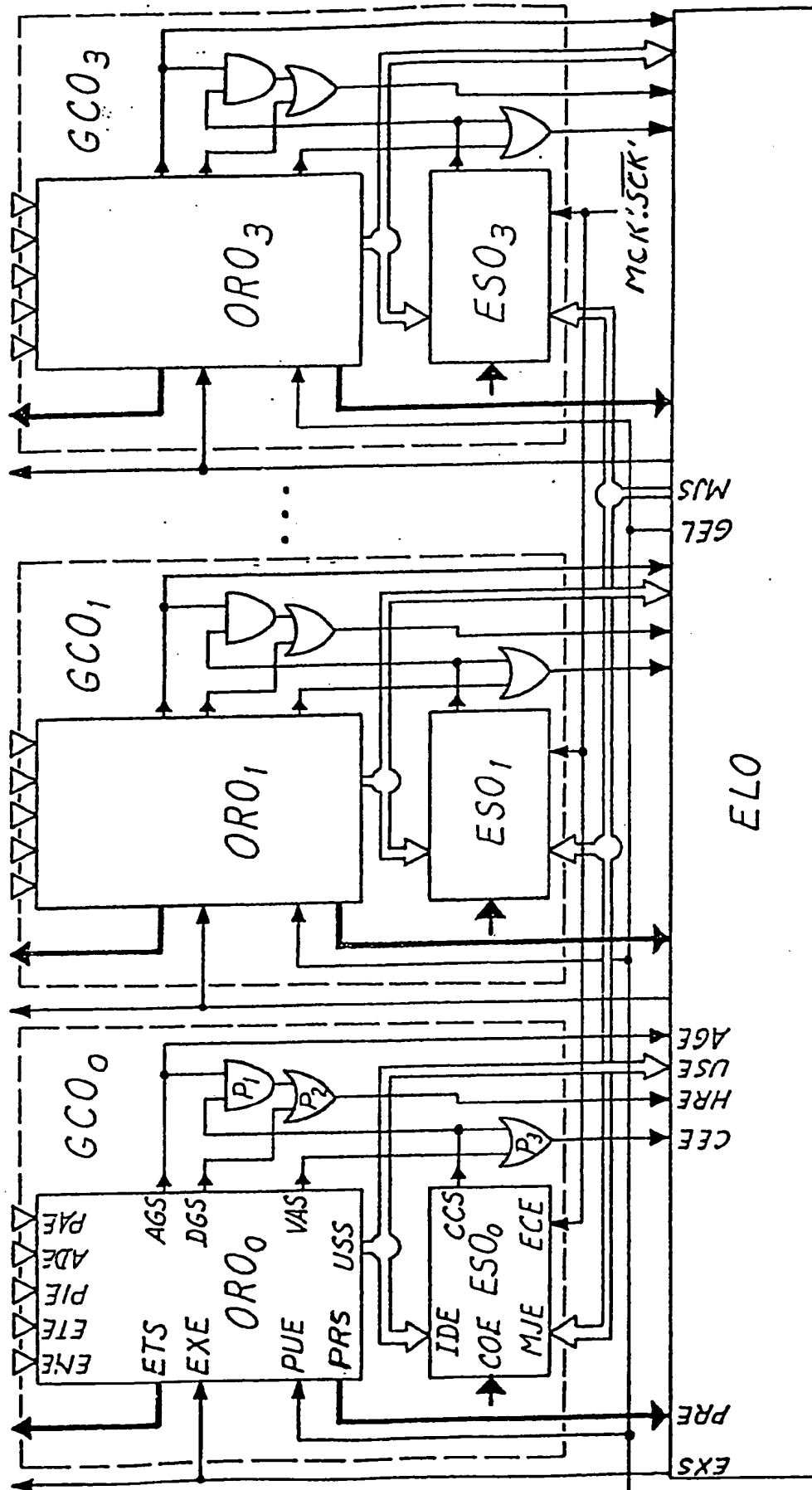


FIG. 3

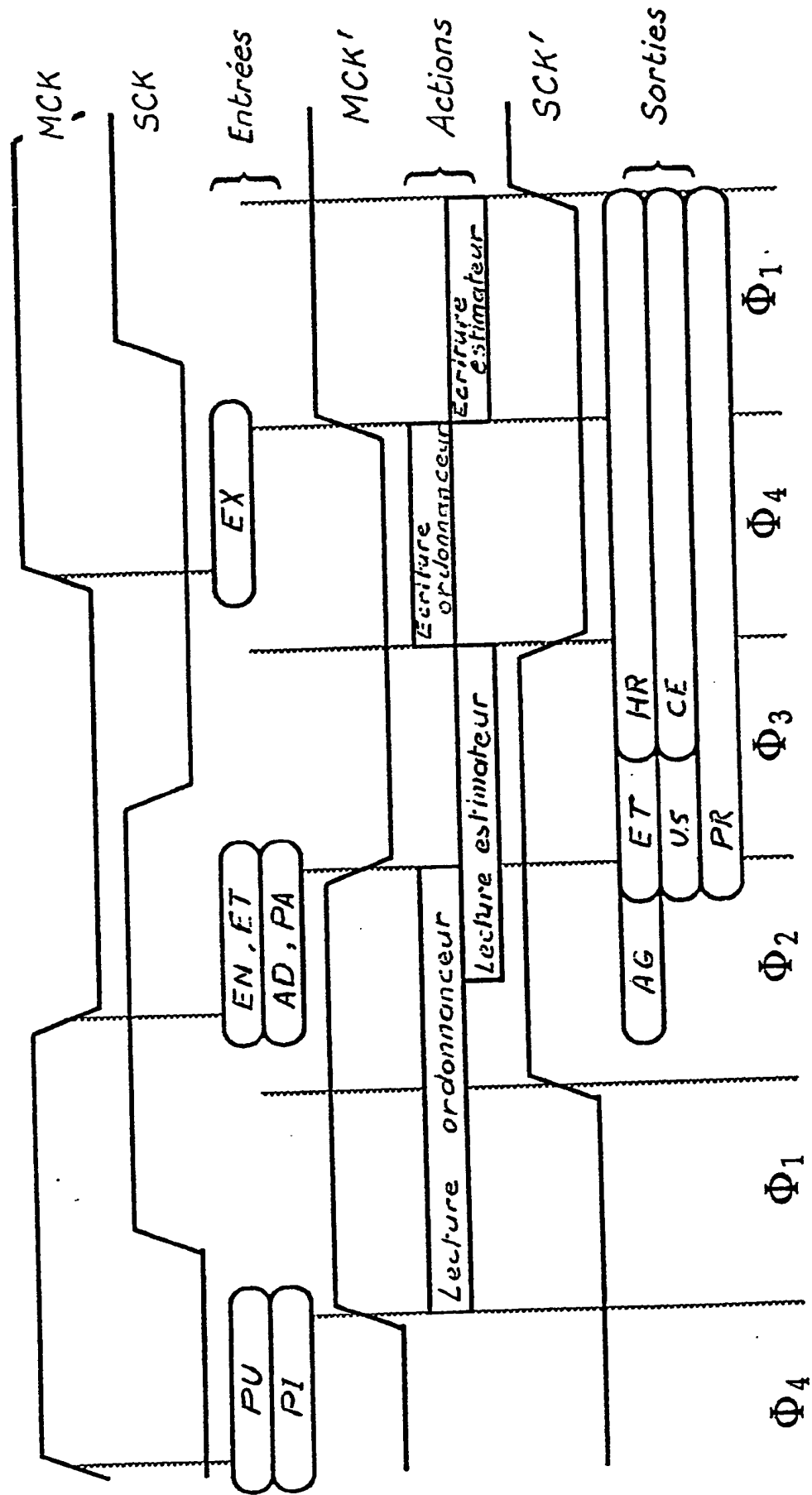


FIG.4

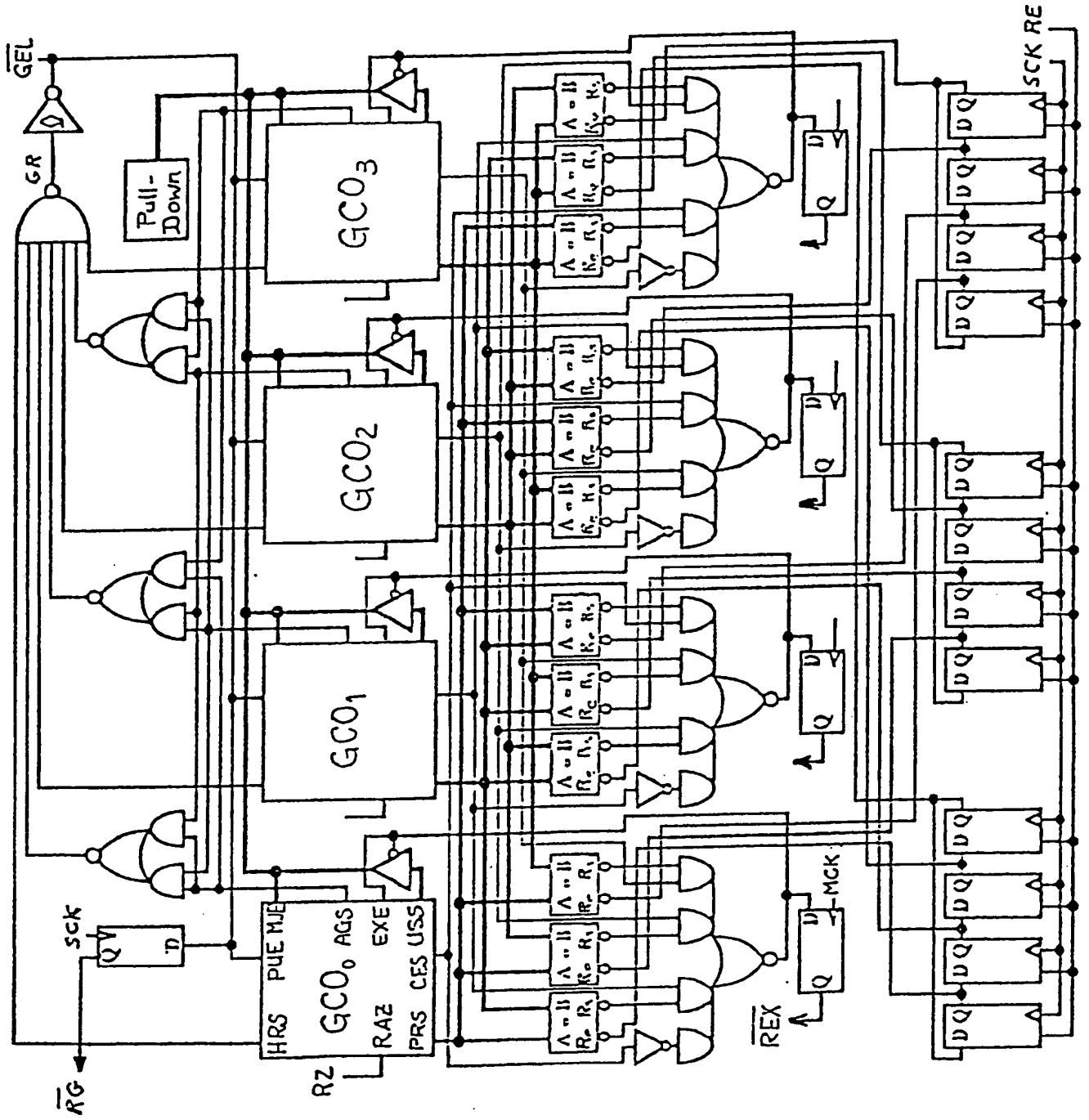


FIG. 5

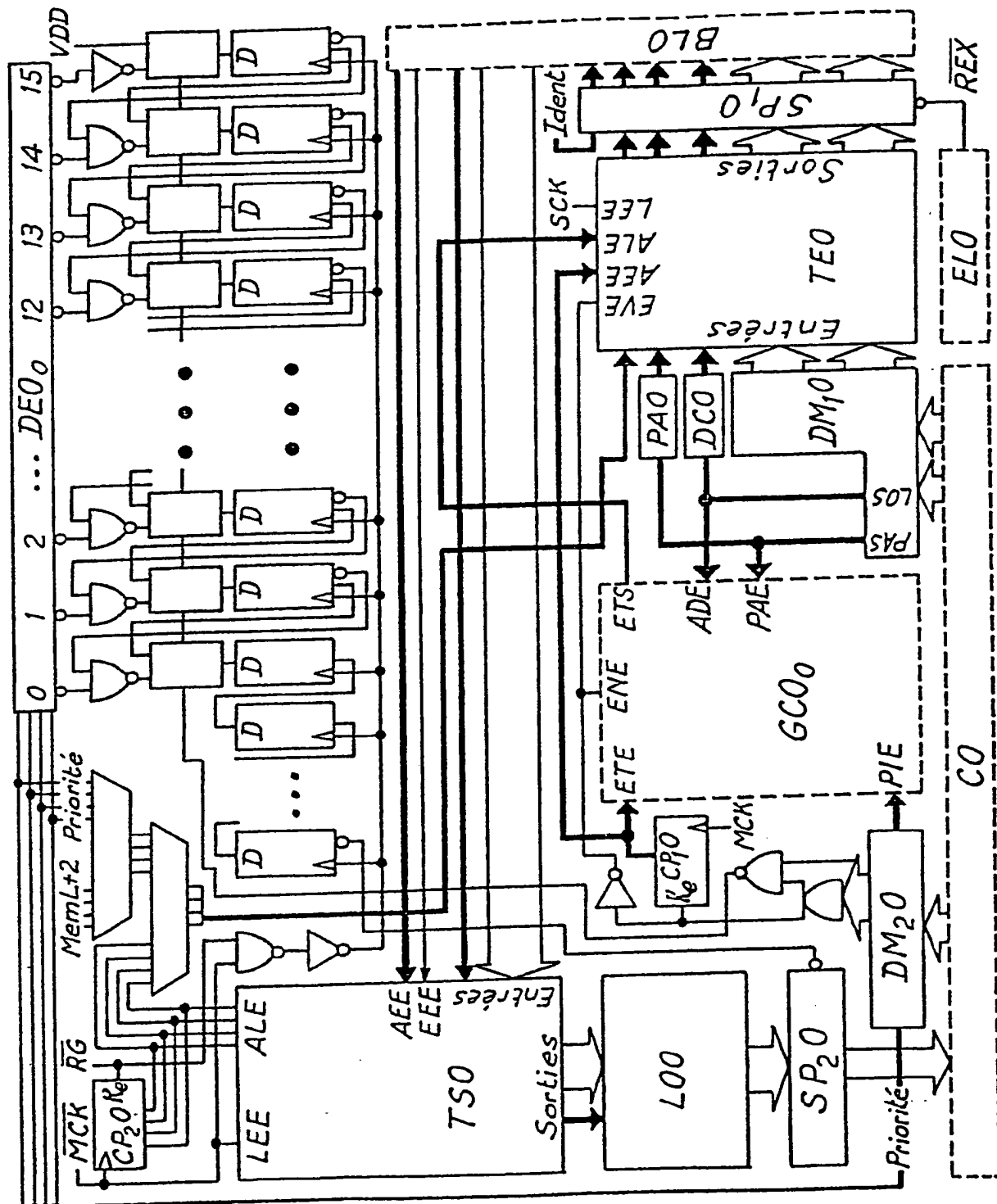


FIG. 6



Office européen  
des brevets

# RAPPORT DE RECHERCHE EUROPEENNE

Numero de la demande

EP 90 40 0148

DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int. Cl.5)
A	EP-A-0 004 775 (A.H. CROXON) * en entier * ---	1,9	G 06 F 13/18
A	IEE PROCEEDINGS SECTIONS A A I vol. 139, no. 4, part E, juillet 1983, pages 116-124, Old Woking, Surrey, GB; E.L. DAGLESS et al.: "Shared memories in the CYBA-M multimicroprocessor"* en entier * ---	1,9	
A	EP-A-0 242 882 (HITACHI LTD.) * en entier * ---	1,9	
A	PATENT ABSTRACTS OF JAPAN vol. 6, no. 88 (P-118)(966), 26 mai 1982; & JP - A - 57 025054 (NIPPON DENKI K.K.) 09.02.1982 * en entier * -----	1,9	
Le présent rapport a été établi pour toutes les revendications			
Lien de la recherche BERLIN		Date d'achèvement de la recherche 28-03-1990	Examinateur DURAND J.
CATEGORIE DES DOCUMENTS CITES X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire		T : théorie ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons ----- & : membre de la même famille, document correspondant	

EPO FORM 1503 03.82 (P0402)

**THIS PAGE BLANK (USPTO)**